The *Chandra*-COSMOS survey IV: X-ray spectra of the bright sample

G. Lanzuisi[1,2,3], F. Civano[2], M. Elvis[2], M. Salvato[1], G. Hasinger[4], C. Vignali[5,6],
G. Zamorani[6], T. Aldcroft[2], M. Brusa[5,6,1], A. Comastri[6], F. Fiore[7],
A. Fruscione[2], R. Gilli[6], L. C. Ho[8], V. Mainieri[9], A. Merloni[1], A. Siemiginowska[2]

[1]*Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 85748 Garching, Germany*
[2]*Harvard-Smithsonian Center for Astrophysics, 60 Garden St. Cambridge, MA 02138.*
[3]*Technische Universität München, Fakultät für Physik, James-Franck-Str. 1 85748 Garching, Germany*
[4]*Institute for Astronomy, 2680 Woodlawn Drive Honolulu, HI 96822-1839 USA*
[5]*Dipartimento di Astronomia, Università degli Studi di Bologna, Via Ranzani 1, I–40127 Bologna, Italy*
[6]*INAF–Osservatorio Astronomico di Bologna, Via Ranzani 1, I–40127 Bologna, Italy*
[7]*Osservatorio Astronomico di Roma (INAF), Via Frascati 33, I–00040 Monteporzio Catone, Italy*
[8]*The Observatories of the Carnegie Institute for Science, Santa Barbara Street, Pasadena, CA 91101, USA*
[9]*ESO Headquarters Karl-Schwarzschild-Str. 2 85748 Garching bei Müchen Germany*

ABSTRACT

We present the X-ray spectral analysis of the 390 brightest extragalactic sources in the *Chandra*-COSMOS catalog, showing at least 70 net counts in the 0.5-7 keV band. This sample has a 100% completeness in optical-IR identification, with $\sim$75% of the sample having a spectroscopic redshift and $\sim$25% a photometric redshift. Our analysis allows us to accurately determine the intrinsic absorption, the broad band continuum shape (Γ) and intrinsic L_{2-10} distributions, with an accuracy better than 30% on the spectral parameters for 95% of the sample. The sample is equally divided in type-1 (49.7%) and type-2 AGN (48.7%) plus few passive galaxies at low z. We found a significant difference in the distribution of Γ of type-1 and type-2, with small intrinsic dispersion, a weak correlation of Γ with L_{2-10} and a large population (15% of the sample) of high luminosity, highly obscured (QSO2) sources. The distribution of the X-ray/Optical flux ratio (Log(F_X/F_i)) for type-1 is narrow ($0 < X/O < 1$), while type-2 are spread up to $X/O = 2$. The X/O correlates well with the amount of X-ray obscuration. Finally, a small sample of Compton thick candidates and peculiar sources is presented. In the appendix we discuss the comparison between *Chandra* and *XMM–Newton* spectra for 280 sources in common. We found a small systematic difference, with *XMM–Newton* spectra that tend to have softer power-laws and lower obscuration.

Key words: Galaxies: active – Galaxies: high-redshift – Galaxies: nuclei – X-ray: galaxies – X-ray: survey

1 INTRODUCTION

There is now strong observational evidence that the accretion onto super massive black holes (SMBHs) and the formation and evolution of their host galaxies are strongly coupled processes. A narrow correlation between the mass of SMBHs and the properties of their host galaxies is observed (Kormendy & Richstone 1995, Silk & Rees 1998, Magorrian et al. 1998, Ferrarese & Merritt 2000). There is also evidence for a common cosmic "downsizing" in both star formation rate (SFR) and active galactic nuclei (AGN) activity (Cowie et al. 1996; Hasinger et al. 2005; La Franca et al. 2005;

Bongiorno et al. 2007; Silverman et al. 2009; Serjeant et al. 2010). These evidences imply that the evolution of galactic bulges and central BHs should be tightly linked through some sort of physical mechanism or feedback (but see also Jahnke et al. 2011 for a different interpretation). For example, luminous AGN can heat the interstellar matter through winds, shocks and ionizing radiation, inhibiting star-formation and accretion itself (Croton et al. 2006; Li et al. 2007; Hopkins et al. 2008, Cattaneo et al. 2009). To understand this feedback mechanism, and hence the evolution of the BH-Galaxy relation, beyond the local Universe (Merloni et al. 2010, Peng et al. 2010), we need to constrain the overall AGN census in detail and to

determine the distribution of AGN parameters in classes and their evolution in redshift.

X-ray surveys are the most effective way of selecting AGN, given that they are less affected by obscuration, up to Compton thick (CT)[1] regime, and contamination form the host galaxy, that strongly affect optical and IR bands (e.g. Donley et al. 2012). This is especially true for all those classes of 'elusive' AGN that are hard to identify as such: X-ray Bright Optically Normal Galaxies (XBONGs, Comastri et al. 2002, Civano et al. 2007); high-z, obscured AGN in highly starforming galaxies (Alexander et al. 2002, 2005, Daddi et al. 2007, Fiore et al. 2008, 2009, Lanzuisi et al. 2009); or local Low Luminosity AGN (LLAGN, Ho et al. 1995). However, the combination with multi-wavelength spectroscopic and photometric observations is required in order to determine the redshift of the source and the host properties, key information to fully understand how SMBHs interact with their environment.

It is still not clear why, to a first approximation, the spectral energy distribution of all quasars are very similar, over a large range of wavelengths and of physical parameters such as BH mass, accretion rate, luminosity and over 13 Gyr of cosmic time. Whatever processes produce the features we observe, there must be some robust physics that does not depend strongly on any of these parameters.

While high signal to noise (S/N) sources show a variety of complex spectral features - warm absorbers, soft excess, emission and absorption lines, a reflection component (e.g. Risaliti & Elvis 2004 for a review) - for low S/N sources the typical X-ray (0.5-10 keV) spectrum can be modeled as a power-law, modified by intrinsic absorption from the surrounding material (the dusty 'torus' or the host galaxy itself). The X-ray spectroscopy of large samples of AGN, with available adequate multi-wavelength coverage, allows one to study the distribution of at least three fundamental parameters such as the column density N_H, the photon index Γ[2] and the intrinsic X-ray luminosity.

Typically all AGN show very similar photon indexes of the primary power-law, with average value $\Gamma \sim 1.9$ and small dispersion, $\sigma_\Gamma \sim 0.2$ (Nandra & Pounds 1994; Piconcelli et al. 2005; Tozzi et al. 2006; Mainieri et al. 2007). This emission is due to the inverse Compton scattering of the UV photons, emitted from the disk, by the hot electrons in the corona surrounding the disk (Haardt & Maraschi 1999; Siemiginowska 2007). This process tends to produce similar slopes, independently of the expected large range of physical parameters of the central engine, such as BH mass, spin or accretion rate. To build a large sample of sources for which these fundamental parameters are known with a good accuracy is not trivial. Even the search for dependencies of the photon index from other observables, such as luminosity, redshift or accretion rate, has not yet produced definitive results (Page et al. 2005; Mateos et al. 2005; Kelly et al. 2007; Young et al. 2009; Sobolewska & Papadakis 2009, Lusso et al. 2012).

X-rays are very effective in probing the distribution of obscuration around the central SMBH, at least in the Compton thin regime, revealing a large fraction of obscured sources both in the local Universe (Risaliti et al. 1999) and at high redshift (Tozzi et al. 2006). This obscuration can be used as an indicator that the host

galaxy is undergoing feedback processes (Hopkins et al. 2008), i.e. the emission from the AGN is interacting with gas and dust in the environment. The fraction of obscured AGN is also depending on the X-ray luminosity (Lawrence & Elvis 1982; Ueda et al. 2003; La Franca et al. 2005, Hasinger et al. 2006, Della Ceca et al. 2008). Such a trend has been observed also in the optical (Simpson 2005) and in the near infrared (Maiolino et al. 2007) and may be linked to the AGN radiative power (coupled with the age): the AGN emission is thought to be able to ionize and expel gas and dust from the nuclear regions, as suggested in the so-called receding torus models (Lawrence 1991; Ballantyne et al. 2006), indicating a breakdown of the standard AGN unification model (Antonucci 1993). The possible evolution of the fraction of obscured sources with redshift is still matter of debate (La Franca et al. 2005; Treister & Urry 2006; Hasinger et al. 2006, Gilli et al. 2007, Della Ceca et al. 2008).

Finally, the X-ray luminosity is the most efficient way to probe the total luminosity due to accretion onto SMBH, because it is less affected by obscuration, with respect to Optical-UV radiation from the accretion disk itself. Together with the BH mass, it allows to derive the accretion rate of the system. Also, the luminosity function is the key ingredient to determine the cosmic history of accretion and the duty cycle of AGN.

We took advantage of the superb dataset provided by the *Chandra* coverage of the COSMOS field (Elvis et al. 2009, Paper I hereafter), to investigate the X-ray emission of a large population of quasars, spanning a wide range of luminosities and redshift, with the aim of studying the distribution of the spectral parameters within the sample and in different optical classes. The high number of sources available with optical/IR identification, especially at high redshift and high luminosities, makes this sample unique among X-ray deep surveys. We limited our analysis to a bright subsample, in order to test our procedure and obtain strong constraints on the spectral parameters.

In the following we present the X-ray spectral analysis of the brightest 390 extragalactic sources in the *Chandra*-COSMOS catalog, with at least 70 net counts in the 0.5-7 keV band. The detection procedure is described in Puccetti et al. (2009, Paper II). This *Chandra*-COSMOS Bright Sample (CCBS hereafter) has 100% completeness in optical-IR identification (Civano et al. 2012, Paper III). The paper is organized as follows: Section 2 summarizes the properties of the COSMOS survey; Section 3 presents the data reduction and sample general properties; Section 4 reports the X–ray fit procedure and results; in Section 5 the fit parameter distributions are discussed, while section 6 summarizes the results. A standard Λ cold dark matter cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_\Lambda = 0.73$ and $\Omega_M = 0.27$ is assumed throughout the paper. Errors are given at 1σ confidence level for one interesting parameter, unless otherwise specified.

2 THE COSMOS SURVEY

The Cosmic Evolution Survey (COSMOS, Scoville et al. 2007a) is a deep and wide extragalactic survey which covers a 2 deg^2 equatorial (10^h, $+02^\circ$) field with imaging by most of the major space-based telescopes (*Hershel*, *Spitzer*, *Hubble Space Telescope*, GALEX, *XMM–Newton* and *Chandra*) and a number of large ground based telescopes (*Subaru*, VLA, ESO-VLT, UKIRT, NOAO, CFHT, and others). Large dedicated ground-based spectroscopy programs in the optical with Magellan/IMACS (Trump et al. 2007), VLT/VIMOS (Lilly et al. 2008, 2009), *Subaru*-FMOS and DEIMOS-*Keck* have been completed or are well underway. The

---

[1] If the X-ray obscuring matter has a column density equal to or larger than the inverse of the Thomson cross-section, $N_H > \sigma_T^{-1} \sim 1.6 \times 10^{24} cm^{-2}$, the source is called Compton thick.

[2] defined as $N(E) \propto E^{-\Gamma}$, where N(E) is the number of photons of energy E.

location of COSMOS will allow all major future facilities (JVLA, ALMA, and SKA) to target this region. Newly awarded follow ups with *Chandra*, *Nu-Star*, JVLA and *Spitzer* will rejuvenate the database in the next years.

This wealth of data has resulted in a 35-band photometric catalog of $\sim 10^6$ objects (Capak et al. 2007) resulting in photo-z's for the galaxy population accurate to $\Delta z/(1+z)<1\%$ (Ilbert et al. 2009) and to $\Delta z/(1+z)\sim 1.5\%$ for the AGN population (Salvato et al. 2009, 2011).

The *Chandra*-COSMOS survey (C-COSMOS) covered the central 0.9 deg^2 region of the COSMOS field, with the ACIS-I CCD imager (Garmire et al. 2003) on board the *Chandra* X-ray Observatory (Weisskopf et al. 2002). The survey used a total of 1.8 Msec of *Chandra* observing time ($\sim$21 days), employing a series of 36 heavily overlapping ACIS-I 50 ksec pointings to give a highly uniform exposure of $\sim$160 ksec over the inner (0.5 deg^2) area and $\sim 70-80$ks in the outer (0.4 deg^2) area. The depth is $\sim 1.9\times 10^{-16}$erg cm^{-2}s^{-1} (0.5-2 keV), three times below the corresponding flux limits for the XMM-COSMOS survey (Hasinger et al. 2007; Cappelluti et al. 2009), making them complementary surveys. The X-ray source catalog comprises 1761 X-ray point sources detected down to a maximum likelihood threshold $detml = 10.8$[3] in at least one band (0.5-2, 2-8, or 0.5-8 keV). The C-COSMOS survey, data analysis and general properties of the detected sample are described in detail in Paper I. Paper II presented the details of the simulations carried out to optimize the detection procedure. The optical and infrared identifications of the C-COSMOS sources have been presented in Paper III.

In the latest *Chandra* call for proposal (Cycle 14), the extension of the coverage of the COSMOS field has been approved (P.I. F. Civano). The proposal aims to complete, with 2.8 Msec of observation, the deep Chandra coverage of the COSMOS area, taking ~ 1.45 deg^2 to 160 ks depth and covering the remaining 0.75 deg^2 field to 100-50 ks depth, more than doubling the number of detected sources.

3 THE C-COSMOS BRIGHT SAMPLE

3.1 Counts extraction

The extraction of source and background counts, and the creation of response matrices, ARF and RMF, for each source, has been performed with *Yaxx*[4]. *Yaxx* is a *Perl* code, developed by T. Aldcroft at the Harvard Smithsonian Center for Astrophysics, which automatically deals with the extraction and fitting of sources detected in large *Chandra* and *XMM–Newton* surveys, such as ChaMPs (Kim et al. 2004) and COSMOS. *Yaxx* makes use of standard *Chandra* CIAO[5] (Fruscione et al. 2006) and *XMM–Newton* Scientific Analysis Software (SAS[6]) tools for the spectra extraction and ARF and RMF creation. CIAO v. 4.1.2 software and CALDB 4.1.2 calibration files were used for the extraction.

The extraction of counts of all the sources listed in an input catalog, is performed in each observation separately, and then the counts are merged together. Given the C-COSMOS tiling, any individual source can be detected up to four times. The radius of the

source extraction regions take into account the degradation of the PSF with off-axis angle, thus it can change significantly from one observation to the other for the same source, in a range of $2''$-$10''$. Each radius is chosen in order to optimize the signal to noise ratio (S/N) in each observation. The background extraction annuli have an inner radius slightly larger than the source radius, and a fixed outer radius of $30''$. In most cases, the extraction regions had to be masked to avoid CCD gaps, field of view (FOV) edges and other sources. These regions were excluded from the extraction areas.

The response matrices, ARF and RMF, are computed with standard CIAO tools in *Yaxx*. The source and background spectra are then merged together using the *mathpha* tool , and the total ARF and RMF are computed with *addarf* and *addrmf* tools, within the FTOOLS package (Blackburn 1995). The scaling factor of source and background regions changes in each observation, and this is taken into account when producing the final merged events. We developed a procedure in order to ensure that the scaling factor of the final merged event is equal to the mean of the scaling factors of the different observations, weighted for the exposure times (i.e. the contribution to the total number of counts, assuming constant flux). This procedure was performed for all the 1761 sources in the catalog, and the resulting distribution of total net counts defined in this way was used to select our sample.

Among the 1761 sources in the C-COSMOS catalog, we selected sources with enough net counts to allow a spectral fit that could simultaneously determine the power law photon index Γ, normalization, and the intrinsic column density N_H of each source with reasonable uncertainties and minimum correlation. We found that selecting sources with more than 70 net counts in the 0.5-7 keV band and using the Cash statistic (Cash 1979), we were able to determine the spectral parameters with an accuracy of at least 30% in most cases (see Sec. 4.2). The initial resulting *Chandra*-COSMOS bright sample (CCBS) has 405 sources.

Fig. 1 (left panel) shows the distribution of 0.5-7 keV net counts for all the 1761 sources in the catalog, and the threshold of 70 counts chosen to select the CCBS (23% of the total sample).

3.2 X-ray colors

In order to evaluate how the X-ray properties of the sources in the CCBS relate to those of the entire C-COSMOS sample, we computed two X-ray colors (hardness ratios, HR) for all the 1761 sources. The hardness ratios are defined as HR1=(M-S)/(M+S) and HR2=(H-M)/(H+M), where S, M and H are the net source counts in the 0.5-2, 2-4 and 4-7 keV bands respectively. Fig. 1 (right panel) shows the X-ray color-color diagram for the entire *Chandra* sample (grey points). For clarity only sources with errors on HR1 and HR2 smaller than 0.5 are plotted (1168 sources out of 1761). The 405 sources in the CCBS are marked with red circles.

Dashed and dotted lines show the expected HRs for simple power-law spectra with intrinsic absorption (in the observer frame) of log(N_H)= 20, 21, 22, 22.5, 23 cm^{-2} (from bottom to top) and photon index Γ= 3, 2, 1, 0 (from left to right) respectively. The right and lower projections show the distribution of HR1 and HR2 for the total sample (grey) and the CCBS (red). The sources in the CCBS reproduce the distribution of X-ray colors of the total sample in the range of HR1 and HR2 = $-0.7-0$, where the majority of the sources lie, but CCBS is biased toward sources with softer spectra (negative HRs) with respect to the total sample. The CCBS excludes most of the outliers, those with HR1 and HR2 > 0, i.e. sources with flat or inverted spectra. Almost all the sources not detected in one of the bands (HR1 or HR2 = ± 1) are also excluded.
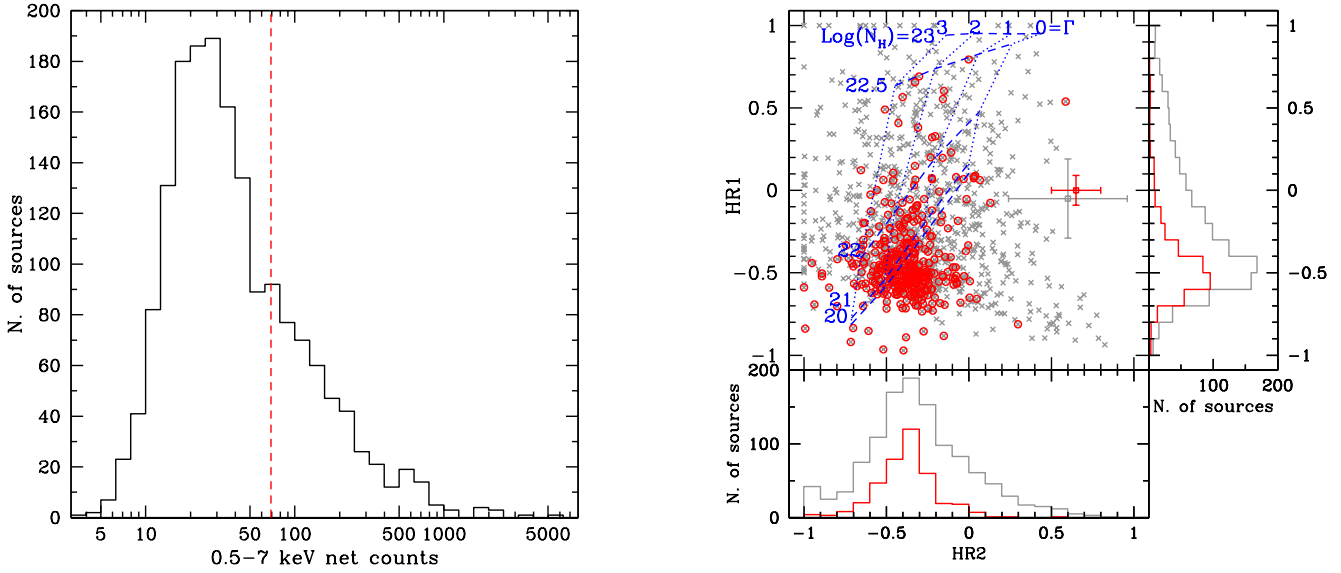

Figure 1. *Left panel:* Distribution of 0.5-7 keV net counts of the total *Chandra* source catalog (1761 sources). The red dashed line marks the threshold we chose in the selection of the 405 sources in the CCBS (i.e. > 70 counts). *Right panel:* X-ray color color diagram for the total *Chandra* catalog (grey points). For clarity only the 1168 sources with errors on HR1 and HR2 smaller than 0.5 are plotted. The grey cross represents the average error for this sub-sample. The 405 sources in the CCBS are labeled with a red circle. The red cross represents the average error for the CCBS. Dashed and dotted lines show the expected HR for simple power-law spectra with intrinsic absorption of $\log(N_H)$= 20, 21, 22, 22.5, 23 cm^{-2} (from bottom to top) and photon index Γ= 3, 2, 1, 0 (from left to right), respectively. The right and lower panels show the distribution of HR1 and HR2 for the total sample (grey) and the CCBS (red).

The 11 sources with HR1 < -0.7 and HR2 < -0.1, i.e. sources with extremely soft spectra, are all comprised in the CCBS and are optically identified as stars (see Sec. 3.3). Indeed the CCBS represents just the tip of the iceberg of the complete catalog and a large number of interesting sources such as heavily obscured, CT AGN and sources with peculiar spectra are located below the threshold of 70 net counts. Thus the extension of this study to fainter sources is certainly of great interest, even if beyond the scopes of this paper.

3.3 Optical-IR identification

The sources in the CCBS have a 99% identification rate. Eleven sources (2.7% of the sample) are classified as stars (Wright, Drake and Civano 2010). Four are not identified: a possible counterpart exists but is close to bright stars or galaxies, for which reliable photometry is not possible and there is not an entry in any COS-MOS photometric catalog (see Paper III for details). We excluded these 15 sources from the sample, to focus on identified extragalactic sources, and we will refer in the following only to these 390 extragalactic sources as CCBS. $\sim$75% (295 sources) of the sample have a spectroscopic redshift and the remaining $\sim$25% (95 sources) a photometric redshift from Salvato et al. (2009, 2011) or Ilbert et al. (2009).

The sample has been divided in different classes on the basis of the optical spectral information or the best fit template used in the SED fitting procedure, following the criteria presented in paper III and Brusa et al. (2010) for the identification. The sources are divided as follows:

• 194 (49.7% of CCBS) are classified as type-1 AGN. Of these, 173 have a spectroscopic redshift showing at least one broad (FWHM> 2000 km/s) emission line and are therefore identified as Broad Line AGN (BLAGN). The remaining 21 have a photo-
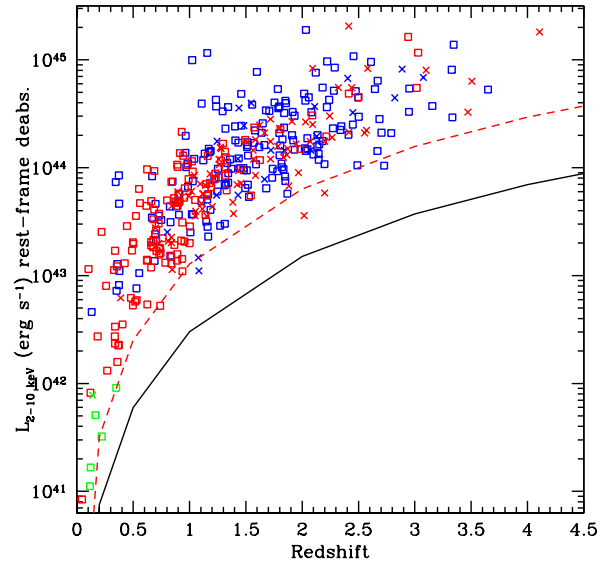

Figure 2. Distribution of the 2-10 keV rest frame luminosity, corrected for absorption, as computed in Sec. 5, as a function of redshift for the 390 sources in the CCBS. In blue are plotted type-1 AGN, in red type-2 AGN and in green galaxies. Open squares (crosses) represent spectroscopic (photometric) redshifts. The solid line represents the sensitivity limit of the whole survey, while the dashed line represent the flux limit estimated for the CCBS.

metric redshift, and the best fit templates are of unobscured quasars of various luminosity with different degrees of contamination by star-forming galaxy templates (from 10 to 90% of the Optical-NIR flux).

Table 1. Distribution of CCBS sources in different classes, and their average redshift and X-ray properties.

Class (1)	Num. (2)	% (3)	z_{min}, z_{max} (4)	% z_{spec} (5)	$\langle z_{spec}\rangle, \langle z_{phot}\rangle$ (6)	$\langle N_{cts}\rangle$ (7)	$\langle F_{2-10keV}\rangle$ (8)	$\langle L_{2-10keV}\rangle$ (9)
Type-1	194	49.7	0.128-3.650	89.2	1.612,1.808	310.5	1.6×10^{-14}	2.4×10^{44}
Type-2	190	48.7	0.045-4.108	61.6	0.919,1.666	166.3	1.4×10^{-14}	1.4×10^{44}
Galaxies	6	1.5	0.116-0.350	83.3	0.194,0.141	91.2	5.2×10^{-15}	4.6×10^{41}

(1) Source class; (2) number of sources; (3) fraction the CCBS sample in each class; (4) minimum and maximum redshift; (5) fraction of spectroscopic redshift; (6) average of the spectroscopic and photometric redshifts; (7) average 0.5-7 keV net counts; (8) average 2-10 keV flux; (9) average 2-10 keV luminosity.

• 190 (48.7% of CCBS) are classified as type-2 AGN. Of these 117 have a spectroscopic redshift and have been identified as non-broad line, i.e. a combination of Narrow Line AGN (NLAGN) and emission or absorption line galaxies. 73 have a photometric redshift, and their best fit SED templates are a combination of obscured AGN templates with different degrees of contamination by passive galaxies, or pure spiral or star-forming galaxy templates. All "passive" galaxies with X-ray luminosity $L_{2-10keV} > 10^{42}$ erg s^{-1} are classified as type 2 AGN (see below).

• 6 (1.5%) are classified as galaxies. Five have a spectroscopic redshift and an optical spectrum typical of passive or pure star-forming galaxies. One has a photometric redshift with a best fit template of passive galaxy. All have $L_{2-10keV} < 10^{42}$ erg s^{-1}.

Table 1 summarizes the distribution of the CCBS sources in different classes. The minimum and maximum redshift, fraction of spectroscopic redshift, average spectroscopic and photometric redshift, average 0.5-7 keV net counts and 2-10 keV flux and luminosity are reported.

Among the 117 sources classified as type 2 on the basis of their optical spectrum and X-ray emission, 32 have a spectrum of NLAGN, according to their position in the diagnostic diagrams presented in Bongiorno et al. (2010). 73 are classified as Emission line galaxies. Of these, 69 have spectral coverage or spectral quality that do not allow to place the sources in the diagnostics diagrams, most of them being at z>1, where these diagrams cannot be constructed, while only 4 are confirmed star forming or LINERs. The remaining 12 are classified as absorption line galaxies (Trump et al. 2009).

Furthermore, among the 73 sources classified as type 2 on the basis of their SED fitting, only 11 have an SED template of type 1.8-2 Seyfert, 16 have a spiral, starforming or ULIRG template from Salvato et al. (2009), while 46 have a template from the library of passive galaxies from Ilbert et al. (2009), (see flow chart in Fig. 6 of Salvato et al. 2011 for details on the template class assignation).

However, all these sources have an X-ray luminosity exceeding the value of $L_{2-10keV} = 10^{42}$ erg s^{-1}. We used this value as a limit to discriminate between star formation and accretion processes. Using the relation in Ranalli et al. (2003) this hard X-ray luminosity correspond to a SFR$\sim$ 200 M$_\odot$/yr, in the ULIRG regime, i.e. they should be quite vigorous (and rare) starbursts. We also stress that 92.5% of the CCBS sample have $L_{2-10keV} > 10^{43}$ erg s^{-1}, that would corresponds to SFR$\sim$ 2000 M$_\odot$/yr, if produced by star formation. There are no sources with this SFR in the COSMOS field (Feruglio et al. 2010) and therefore at the X-ray luminosities of the CCBS we can be sure at least of the presence of an AGN, even if the presence of both components is not excluded. We there-

fore apply an a-posteriori X-ray luminosity cut, and classified all these sources with $L_{2-10keV} > 10^{42}$ erg s^{-1} as type 2 AGN.

These results show that optical diagnostic diagrams and SED model fitting can be insensitive to hybrid objects (e.g. obscured AGNs with enhanced star formation). Indeed 70% of the type-2 with an optical spectrum, and $\sim$ 80% of type-2 with a photo-z from SED fitting, would have been classified as starforming or passive galaxies (or would remain unclassified) on the basis of the optical/IR data only. The coupling of diagnostic diagrams and SED fitting, with the X-ray luminosity cut is indeed a more effective method of separating obscured AGNs from truly nonactive galaxies (see also the discussion in Paper III and Trouille et al. 2011), especially at these flux levels and in the cases of hybrid objects or in all the cases in which the AGN contribution is overwhelmed in optical/IR by the host-galaxy light.

Fig. 2 shows the 2-10 keV rest frame luminosity, corrected for absorption as computed in Sec. 5, as a function of the redshift for the CCBS. Empty squares (crosses) represent sources with spectroscopic (photometric) redshift. Blue, red and green colors represent type-1, type-2 and galaxies, respectively. The sensitivity limit of the total *Chandra*-COSMOS survey is represented by the solid line, while the sensitivity limit estimated for the flux of a source with 70 net counts and a typical AGN spectrum (a simple power law with $\Gamma = 1.9$, $N_H = 1 \times 10^{21}$ cm^{-2}) is reported with a dashed line.

4 SPECTRAL ANALYSIS

4.1 Spectral fit

The spectral fitting was performed using *Sherpa*[7], the CIAO modeling and fitting package (Freeman, Doe and Siemiginowska 2001). We performed a spectral fit based on the Cstat statistics, a modified version of the Cash statistic (Cash 1979), which is well suited for the low counts regime because, in principle, it does not require counts binning. The advantage of Cstat as implemented in the fitting package *Sherpa*, with respect to the Cash statistic, is that the change in Cstat from one model fit to the next (ΔC), is distributed approximately as $\Delta\chi^2$, and it is possible to estimate the goodness of the fit[8], i.e., the observed statistic, divided by the number of degrees of freedom (DOF), should be of order 1 for good fits. However, this is true only when the number of counts in each bin is
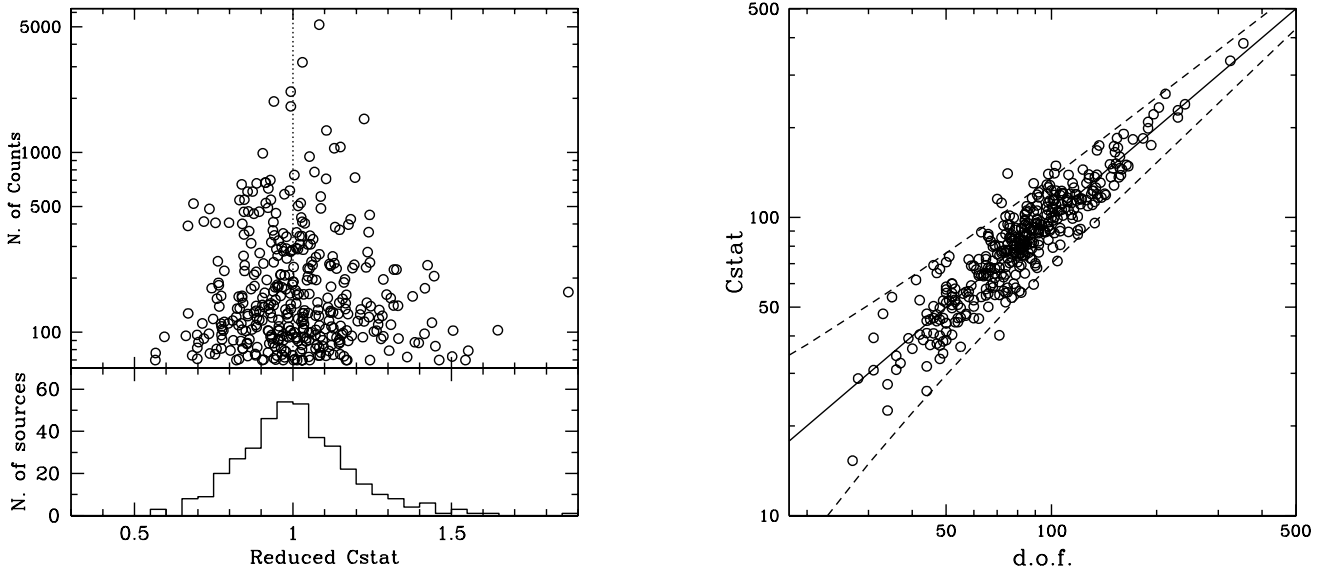
---

[7] http://cxc.harvard.edu/sherpa/index.html
[8] http://cxc.harvard.edu/sherpa/statistics/index.html

Figure 3. *Left panel:* The reduced Cstat (Cstat$_\nu$) as a function of the 0.5-7 keV net counts for the CCBS sample is shown. The bottom panel shows the histogram of the Cstat$_\nu$ for the sample. The dotted line shows the expected value for a good fit of Cstat$_\nu = 1$. *Right panel:* Distribution of Cstat with respect to DOF for each source. The thick line represents the ideal line of Cstat$_\nu = 1$, while the dashed lines show, for a given number of DOF, the value of Cstat above (below) which we expect a 1% probability to find such high (low) value if the model is correct.
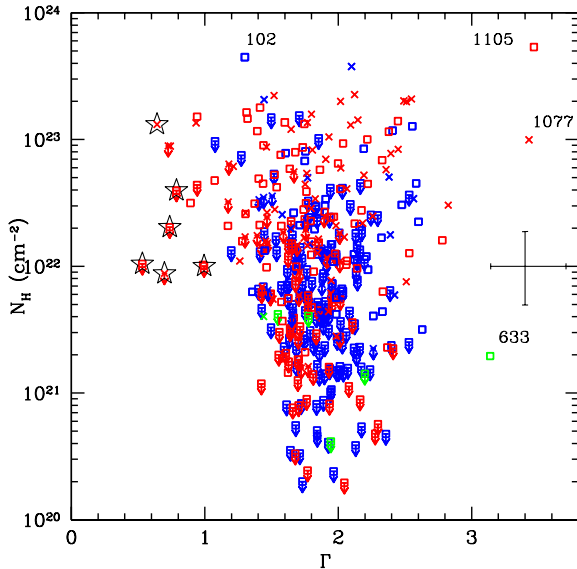

Figure 4. Distribution of Γ and N_H for the CCBS sample. Arrows represent 1σ upper limits in the value of N_H. Symbols as in Fig. 2. Peculiar sources discussed in Sec. 4.3 are labeled with the CID number. Black stars are CT candidates on the basis of their photon index, discussed in Sec. 4.3.5.

$\gg 1$. Therefore, we binned the spectra to 5 counts per bin (see Appendix A for details).

Cstat is a maximum likelihood function and assumes that the error on the counts is purely Poissonian, and so it cannot deal with data that have already been background subtracted. Therefore, one has to model the background, and include this model in the fit. In the *Sherpa* package this is accomplished in one step

by performing a simultaneous fit of the background and the total (source+background) spectra.

The global *Chandra* ACIS-I background between 0.5 and 7 keV is complex, and can be reproduced by two power laws, modified by photoelectric absorption, plus 3 narrow Gaussian lines to reproduce the features at 1.48, 1.74 and 2.16 keV (Al, Si and Au instrumental lines), plus a thermal (MEKAL) component. All these components account for the sum of the particle, cosmic and galactic components of the background (Markevitch et al. 2003, Fiore et al. 2012). The global background shape was recovered from background spectra extracted from a large (8′ radius) regions from different pointings, excluding the astrophysical sources, in order to collect a large number of counts. This allows for a detailed global characterization of the background. The model is then rescaled (for the extraction area) to each local background and fitted, leaving as free parameters the overall normalization and the photoelectric absorption, to account for variations due to different position in the detector, which are typically more prominent in the soft (0.5-2 keV) band. The background spectra for all the CCBS sources have very similar shape and normalization and the background accounts for just $\sim 10\%$ percent of the total flux even at 70 net counts level.

Given the limited number of counts available, the spectral fit of the sources was performed with a simple model: a power law, modified by photoelectric absorption at the source redshift, plus Galactic absorption, fixed to the value in the direction of the field ($N_{H,Gal} = 2.5 \times 10^{20}$ cm^{-2}, Kalberla et al. 2005).

The distribution of the reduced Cstat (Cstat$_\nu$) of the best fits for the 390 sources in the CCBS sample is shown in Fig. 3 (left panel) as a function of the number of 0.5-7 keV net counts. The bottom panel shows the histogram of Cstat$_\nu$ for the sample. The distribution has a peak around 1 as expected, with $\sigma = 0.20$[9].

[9] The distributions are different if the spectra are either unbinned or binned with only 1 or 5 counts per bin. For a detailed discussion on the behavior of
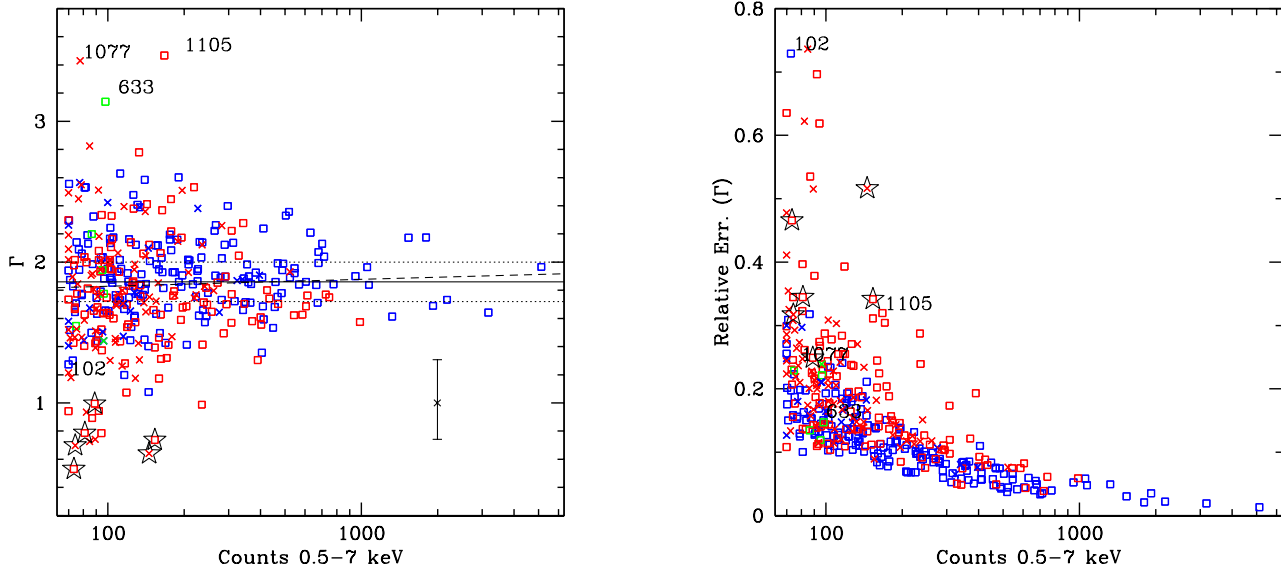
Figure 5. *Left panel:* Distribution of Γ as a function of the 0.5-7 keV net counts for the CCBS sources. Symbols as in Fig. 2. The solid line represents the mean of the distribution, and the dotted lines show the intrinsic standard deviation (see text). The dashed line is the least-squares fit weighted by the errors in Γ. *Right panel:* Distribution of relative errors on Γ with respect to the number of 0.5-7 keV net counts. The accuracy on the parameter determination is better than 30% for 95% of the sources. In both panels peculiar sources and CT candidates are labeled as in figure 4.

Fig. 3 (right panel) shows the distribution of Cstat with respect to DOF for each source. The solid line represents the ideal line for a good fit (Cstat$_\nu = 1$), while the dashed lines show, for any given number of DOF, the Cstat value above (below) which we expect a 1% probability (one-sided) to find such high (low) value if the model is correct. The quality of the fit is very good, given that almost all (383/390) the sources in our sample are distributed between the two dashed lines, i.e. the simple model used is in general able to satisfactorily reproduce the spectra: We find only 7 sources (1.7% of the sample) above the upper dashed line, and 3 sources (0.7%) below the lower one. In 3 cases these sources have actually more complex spectra, requiring additional components in the model to be reproduced (see below). The others are instead more (or less) noisy (either in the source or in the background spectra) than average. As an example, the source with the highest value of Cstat with respect to the DOF is the source with *Chandra* ID (CID)-1105, which has Cstat$_\nu$=1.87. This source has a complex spectrum, showing both a strong obscuration and an extra soft component. This peculiar source, together with others are discussed in Sec. 4.3.

4.2 Modeling results

Fig. 4 shows the distribution of the two main spectral parameters, Γ and N_H, for the CCBS sample. There is no significant correlation between Γ and N_H in the sample [10] even if there can be in individual sources with low statistics. The photon index is distributed around the typical value for AGN ($\langle\Gamma\rangle = 1.84$) with observed dispersion $\sigma_{obs} = 0.39$. This dispersion is the result of the convo-

lution of the intrinsic distribution with the error distribution. Using the maximum likelihood method of Maccacaro et al. (1988), and assuming that both distributions are Gaussian, we estimated an intrinsic dispersion $\sigma_{int} = 0.22$, a very small dispersion, indicating high homogeneity in the sample spectra. The sources with low obscuration ($N_H < 10^{22}$ cm^{-2}) show a smaller dispersion in Γ ($\sigma_{obs} = 0.29$, $\sigma_{int} = 0.09$), while the spread and the errors in Γ increase strongly with increasing N_H. This is due to the fact that a large amount of obscuration makes it more difficult to retrieve the intrinsic value of the photon index.

From Fig. 4 it is possible to identify few sources with peculiar spectra, namely three extremely soft sources, with $\Gamma > 3$, (between 2 and 3σ from the mean Γ value), and six flat sources, showing a photon index significantly flatter (at 3σ) than the mean Γ values, for which a reflection dominated model may be a better description of their X-ray spectra. All these sources are discussed in Sec. 4.3 (we also discuss there source CID-102, showing the highest value of N_H in the sample). We excluded these sources when computing the distribution of Γ. After excluding these outliers, the results are $\langle\Gamma\rangle = 1.86$ with dispersions $\sigma_{obs} = 0.34$ and $\sigma_{int} = 0.14$, i.e. the intrinsic dispersion is reduced of almost 50%.

Fig. 5 (left panel) shows the distribution of photon index Γ as a function of the number of source counts for the CCBS sample (symbols as in Fig. 2). The mean value $\langle\Gamma\rangle = 1.86$ and the intrinsic standard deviation $\sigma_{int} = 0.14$ are reported with solid and dotted lines respectively. A least-squares fit, weighted by the errors in Γ, gives an almost flat slope (dashed line). As expected, for lower number of counts, the spread in the values of Γ increases, along with the errors. The mean value of the distribution is close to the average values found for large samples of quasars in the 2-10 keV band (e.g. Reeves & Turner 2000; Piconcelli et al. 2005, Tozzi et al. 2006). Fig. 5 (right panel) shows the distribution of relative errors on Γ with respect to the number of source counts. As can be seen, the photon index can be determined with an uncertainty smaller than 30% in 95% of the sample.

---

the Cstat minimization technique under different binning conditions, and a comparison with χ^2 statistic see Appendix A.

[10] We test for correlations using the Kendall correlation test (probability P=0.76 of no correlation) available in ASURV, a survival statistics package (Lavalley et al. 1992).

Table 2. Best fit X-ray spectral parameters for the CCBS.

CID	z	Type	Flag[a]	N_H cm^{-22}	Γ	$F_{0.5-2keV}$ cgs	$F_{2-10keV}$ cgs	$L_{2-10keV}$ cgs	Cstat	DOF	Counts
20	1.156	Type-2	SP	$5.77^{+2.86}_{-2.41}$	$1.86^{+0.46}_{-0.42}$	1.23×10^{-15}	6.73×-15	5.49×10^{43}	84.03	98	125.9
21	1.850	Type-1	SP	< 0.23	$1.59^{+0.09}_{-0.09}$	7.49×10^{-15}	1.71×10^{-14}	4.18×10^{44}	95.66	133	412.1

[a] SP (PH) for spectroscopic (photometric) redshifts.

Table 2 summarizes the results from the best fit of the X-ray spectra for each source[11]. We report the CID number, redshift, classification, flag for spectroscopic or photometric classification, N_H and Γ with relative errors, 0.5-2 keV flux, 2-10 keV flux, 2-10 keV luminosity, Cstat value, DOF value, number of counts.

4.3 Peculiar sources

In the following we discuss a small sample of peculiar sources. Four sources have peculiar or complex X-ray spectra[12], while 6 sources have very flat, possibly reflection dominated spectra. The four sources with peculiar spectra are labeled in Fig. 4 and 5 with their CID number, while CT candidates are labeled with stars.

4.3.1 CID 1105

This source at redshift z=0.219 shows the highest value of Cstat with respect to DOF (Cstat$_\nu$=1.87). The source also shows the second highest value of N_H in the sample, and a steep power-law. As can be seen in Fig. 6 (top left panel), the source has a complex spectrum, showing a strongly obscured primary power-law, and an extra soft component superimposed to it. The soft component can be modeled either with a thermal component (kT= $0.20^{+0.09}_{-0.05}$ keV at the source redshift), or with a very soft, second power-law ($\Gamma = 3.8^{+1.4}_{-1.2}$). Both fits with an additional components are significant improvements with respect to the simple absorbed power law fit, having a Cstat$_\nu$=0.99 and 0.98, respectively. The F-test performed to check if the addition of an extra component is required, gives probabilities of $P = 4.23 \times 10^{-13}$ and $P = 1.53 \times 10^{-11}$ respectively, i.e. the extra component is required by the data at very high significance. Given the quality of the data available, it is not possible to establish which one of the two models is preferred by the fit, even if the use of a thermal component results in a slightly better fit. Furthermore, the optical *Hubble*-ACS image shows a bright, extended elliptical galaxy. This suggests that the source is an obscured AGN, with the soft X-ray component coming from thermal emission from the intra-galactic gas, with possible contribution from the AGN scattered component. The results of both fits with additional soft components are consistent ($\Gamma = 2.42^{+0.76}_{-0.47}$, $N_H = 44.0^{+20.3}_{-15.1}$ cm^{-2}, $L_{2-10keV} = 3.7 \times 10^{43}$ erg s^{-1}) and will be adopted in the rest of the discussion and in Table 2.

4.3.2 CID-633

This is the only one source in the CCBS which shows a steep spectrum ($\Gamma = 3.1^{+0.49}_{-0.41}$) and a low value of N_H. The *Chandra* spectrum of the source, which is optically classified as a non active galaxy at z=0.35, is shown in Fig. 6 (top right panel) and can be reproduced with a thermal emission with temperature $kT = 0.35^{+0.06}_{-0.05}$ keV at the source redshift, compatible with emission from hot gas in the galaxy. This is also consistent with a relatively low X-ray luminosity of $L_{2-10keV} \sim 10^{42}$ erg s^{-1}. The ACS image shows that this source has a neighbor, with the two nuclei being at 3.15″. The X-ray emission corresponds to the central region of CID-633. We conclude that it is compatible with emission from a recent star formation event in the central region of CID-633, likely triggered by the interaction with the neighbor galaxy. This source is excluded from the following discussion on the power-law fit parameter distribution.

4.3.3 CID-1077

This source at z=2.017 has as best fit SED a galaxy template from Ilbert et al. 2009. It has a 2-10 keV luminosity of 3×10^{43} erg cm^{-2} s^{-1}. The *Chandra* spectrum of the source (Fig. 6, bottom left panel) shows a high column density (of the order of 10^{23} cm^{-2}) and a steep spectrum ($\Gamma = 3.42^{+0.90}_{-0.72}$). This combination of spectral parameters is quite unusual, and can be due to a thermal spectrum, like in the case of CID-633, if the photometric redshift for this source were incorrect (few % of outliers are expected). However, from the ACS image, it appears that the optical counterpart is very faint, compatible with the high redshift scenario. In this case it is hard to explain the steep power-law characterizing the *Chandra* spectrum of this source.

4.3.4 CID-102

This source has the highest value of N_H in the CCBS sample ($N_H= 4.4^{+3.6}_{-2.4} \times 10^{23}$ cm^{-2}) and a flat spectrum ($\Gamma = 1.30^{+1.04}_{-0.85}$), but is optically classified as a broad line AGN at z=1.84. However, the X-ray emission of this source is just above our selection limit. The X-ray spectrum (Fig. 6 bottom right panel) has 70.2 net counts (0.5-7 keV) and shows a hint of a second soft component below 1 keV. As a result, the source shows one of the highest values of reduced Cstat (Cstat$_\nu$=1.36) and high values of relative errors (the source is the second from the top in Fig. 5 right panel). Adding a soft component (either a second power-law or a black body component) to the model produces a significant improvement in the fit (Cstat$_\nu \sim 1.15$ in both cases), giving similar results on N_H and Γ of the previous model, but with smaller errors. We will use these results in the following discussion. However, given the quality of the

[11] The full table is available online at *https://hea-www.cfa.harvard.edu/~fcivano/C_COSMOS_X-ray_spectral_analysis_CCBS.fits.*
[12] Another source in the CCBS with peculiar features around the Fe $K\alpha$ line, namely CID-42, is discussed in Civano et al. (2010).
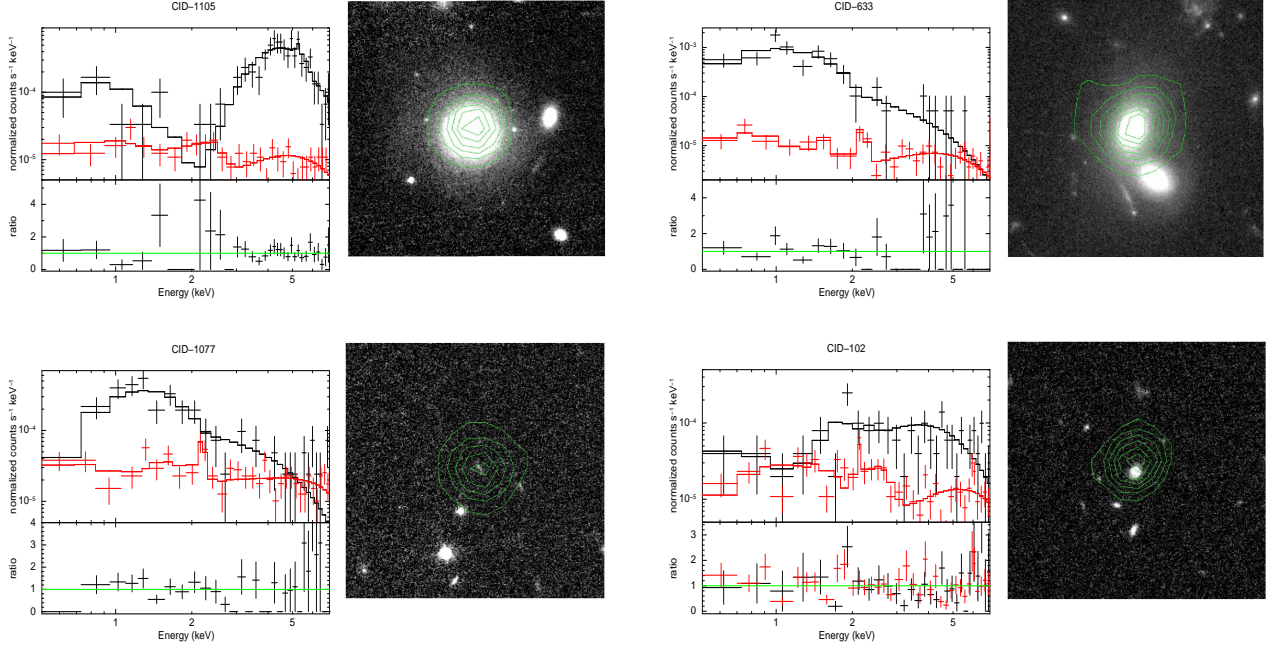
Figure 6. *Chandra* 0.5-7 keV spectra (left panel) and *Hubble*-ACS image (right panel) of four peculiar sources in the CCBS: CID-1105, 633, 1077 and 120, from top left to bottom right. The total spectrum is shown in black, and the background spectrum in red, each with its best fit model. The counts are binned to a minimum significant detection of 3σ (and a maximum of 15 counts per bin) for plotting purposes only. The lower part of each left panel shows the model to data ratio for the total spectra only. In the right panels the *Chandra* ACIS-I contours (in green) are superimposed to the ACS images.
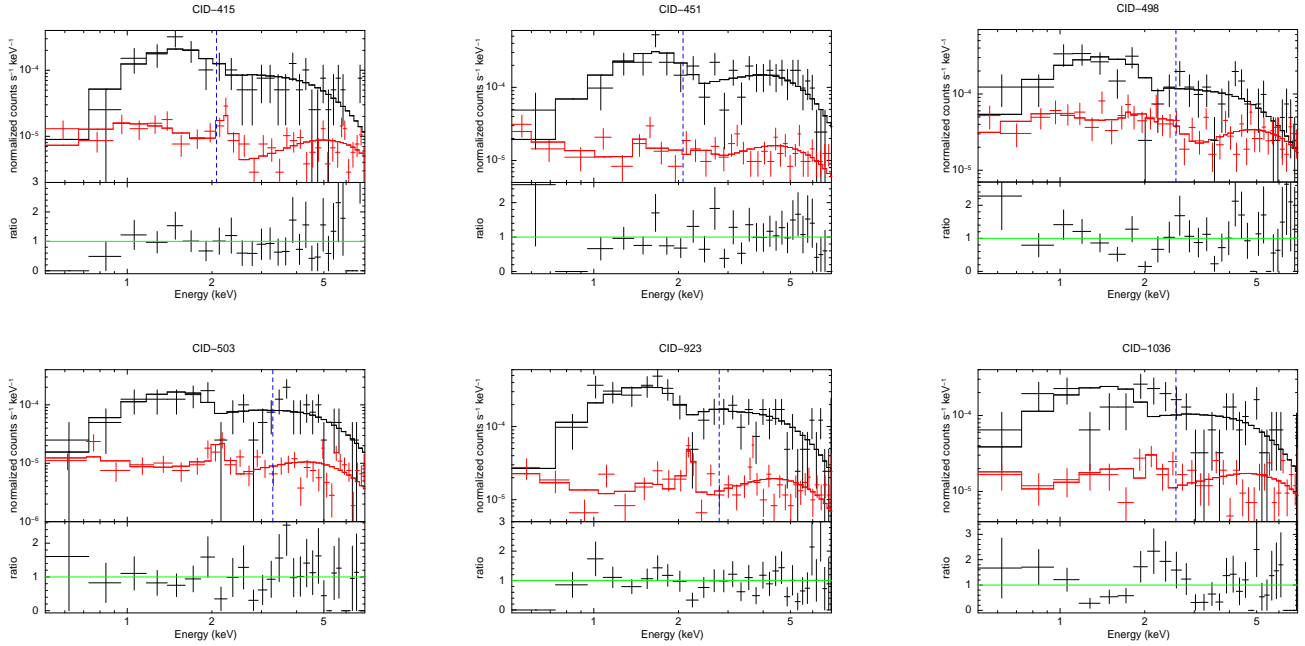
Figure 7. *Chandra* 0.5-7 keV spectra of the CT candidates in the CCBS: CID-415, 451, 498, 503, 923 and 1036 from left to right and top to bottom. The total spectrum is shown in black, and the background spectrum in red, each with its best fit model. The counts are binned to a minimum significant detection of 3σ (and a maximum of 15 counts per bin) for plotting purposes only. The lower part of each panel shows the model to data ratio for the total spectra. The blue dashed line represents for each source the location of the expected Fe Kα emission line at 6.4 keV.

spectrum, we cannot draw any firm conclusions for this peculiar source.

4.3.5 Compton Thick candidates

There is a small population of sources with very flat spectra, for which a reflection dominated, CT model may be a better description of the extremely flat spectrum. The "smoking gun" signature in the spectra of these highly obscured AGN would be a strong
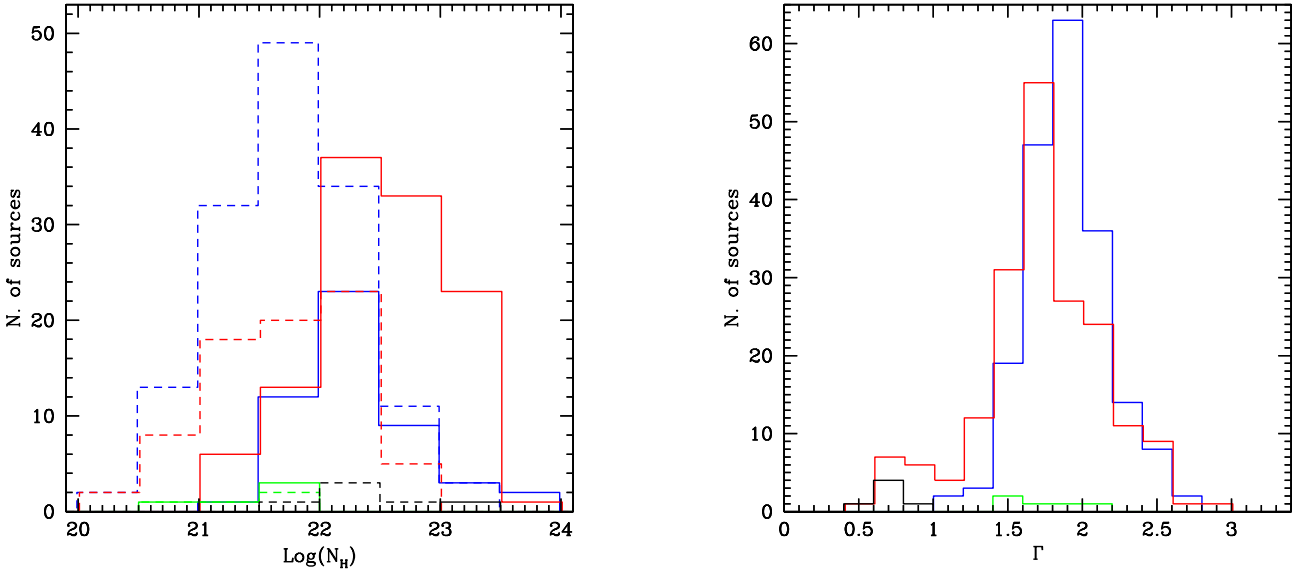
Figure 8. *Left panel:* Distribution of column density N_H. In blue are represented type-1 AGN, in red type-2 AGN, in green sources classified as galaxies, in black CT candidates. The solid (shaded) histograms represent sources with a detection (an upper limit) of N_H. *Right panel:* Distribution of photon index Γ. Colors as in left panel.

(EW~ 1 keV) Fe Kα fluorescent emission line at 6.4 keV (George & Fabian 1991; Matt, Perola, & Piro 1991). However, almost all of these sources have less than 100-200 counts (Fig. 5 left), making difficult to detect even such a strong line.

The spectra of sources which show a photon index significantly flatter than the average Γ, namely CID-415, 451, 498, 503, 923 and 1036, are shown in Fig. 7 from top left to bottom right. The blue dashed line represent for each source the location of the expected Fe Kα emission line at 6.4 keV. A very flat power law, in some case with the addiction of significant (but always $< 3 \times 10^{23}$ cm^{-2}) obscuration, can reproduce the shape of the spectra of these hard sources. However, a completely reflected power law (i.e. *xspexrav* model in Sherpa, based on the *pexrav* model of XSPEC, with pure reflection) can give similar results but, given the low data quality, it is not possible to determine the preferred model on the basis of the fit results. A strong ($EW \sim 1$ keV) Fe Kα line at 6.4 keV is expected as part of the reflected emission, but no sign of such signature is present in four of the observed spectra, with the possible exception of sources CID-503 and CID-1036. The upper limit on the equivalent width (EW) of the emission line is in the range 0.09-0.38 keV, depending on the quality of the data, with average value $\langle EW \rangle = 0.20$ keV.

The source CID-503 shows some residuals at energies just above 6.4 keV, possibly indicating the presence of an Fe Kα emission line. If the energy and normalization of the line are left free to vary (with fixed $\sigma = 0.01$ keV) we obtain an emission line at energy $E_{line} = 6.93^{+0.11}_{-0.11}$ keV. The equivalent width is EW$= 0.65^{+0.25}_{-0.21}$ keV. The energy of the line is compatible, within 2σ, with a ionized Fe line at 6.7 keV, that appears to increase in importance in high redshift type-2 AGN (Iwasawa et al. 2012).

For source CID-1036 the best fit is obtained if also the width of the line is left free to vary: a broad ($\sigma = 0.3$ keV) emission line at energy $E_{line} = 5.85^{+0.58}_{-0.52}$ keV, but compatible with 6.4 keV within one σ, is required. We underline that the photometric redshift of the source ($z = 1.497$) has an error of ± 0.02 at 1σ. The

EW is compatible to the one expected from a reflection dominated spectrum (EW$= 1.20^{0.93}_{-0.65}$ keV).

We also computed the expected number of CT in our sample, according to the most recent predictions of X-ray background synthesis models (Gilli et al. 2007, Treister et al. 2009). Using the *pexrav* model of pure reflection we translated the 70 counts (0.5-7 keV) cut into a flux limit of $\sim 1.1 \times 10^{-14}$ erg cm^{-2} s^{-1} in the 2-10 keV band (for a nominal exposure time of 160 ks). The Gilli et al. model[13] predict ~ 4 CT sources per deg^{-2} above this flux limit. This, folded with the C-COSMOS area-sensitivity curve, translates into $2-3$ CT sources expected. Therefore the lack of confirmed CT sources in our sample is not in disagreement with predictions from these background synthesis models, and is consequence of the flux limit of our sample.

5 FIT PARAMETER DISTRIBUTIONS

Having verified that our results are statistically robust (see also Appendix A), we can discuss the distribution of the derived spectral parameters. In Fig. 8 we show the distributions of two spectral parameters N_H and Γ (left and right panels, respectively). In both plots colors identify different classes of sources, blue for type-1, red for type-2, green for galaxies and black for the sample of CT candidates. In the left panel, the solid (shaded) histograms represent detections (upper limits) of N_H.

5.1 N_H, intrinsic absorption column density

For type-1 AGNs, the distribution of intrinsic obscuration has a global peak in the bin between $3 - 10 \times 10^{21}$ cm^{-2}, but is strongly dominated by upper limits (75% of the sources). The fraction of

---

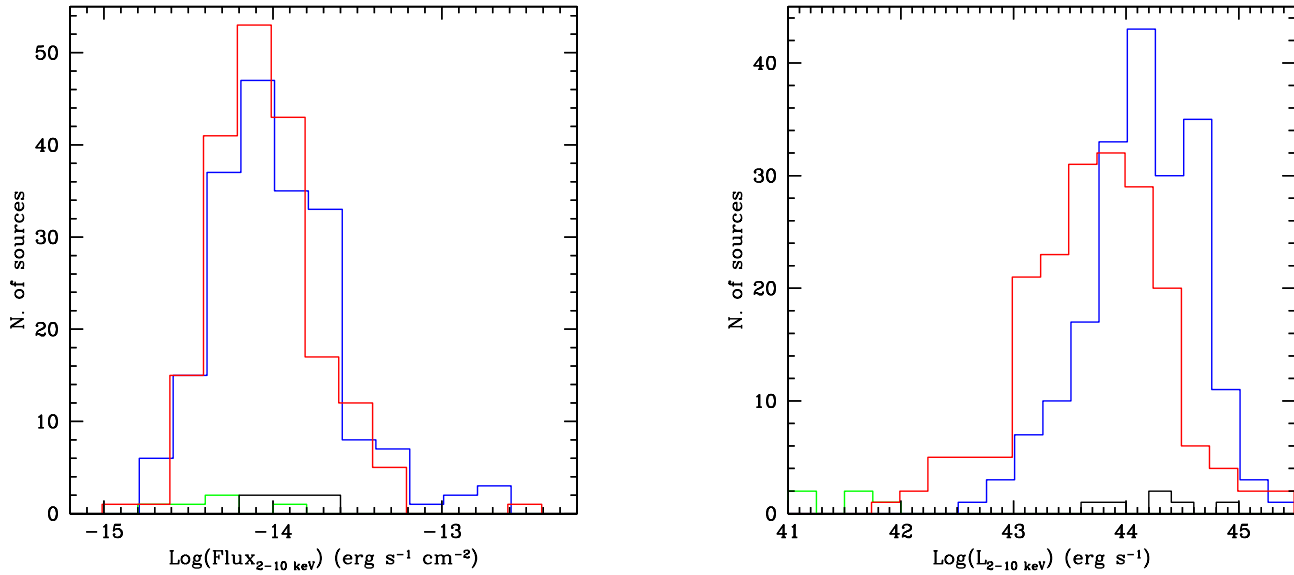[13] http://www.bo.astro.it/ gilli/counts.html

Figure 9. Distribution of 2-10 keV observed flux (left) and 2-10 keV intrinsic luminosity, corrected for absorption (right). Colors as in previous figures.

sources that can be considered obscured ($N_H > 10^{22}$ cm^{-2}), assuming a Gaussian distribution for the errors, and computing for each source the fraction of the Gaussian above 10^{22} cm^{-2}, is 17% (33 sources). This fraction is slightly larger of what found in shallower surveys (Brusa et al. 2003; Silverman et al. 2005; Mateos et al. 2010; Corral et al. 2011; Scott et al. 2011), and similar to what found for deeper surveys, like the 1 Msec. observation in the CDFS (Tozzi et al. 2006).

For type-2 AGN instead, the situation is inverted. The distribution is peaked above 10^{22} cm^{-2}, and 60% of the sources have a detection of intrinsic column density. However, the fraction of type-2 that can be considered obscured (computed as for type-1s), is only 50%. This is a low value for this class of sources, compared with previous results. This may be due to the fact that our sample is biased against low counts sources, and hence in favor of less obscured ones (e.g. at $F_{2-10keV} \sim 5 \times 10^{-15}$ a source with $N_H = 5 \times 10^{22}$ cm^{-2} is excluded by our counts based selection, while an unobscured source is not). Furthermore $\sim 35\%$ of non obscured type-2s are classified on the basis of their SED fitting, typically as passive galaxies, and not from a proper optical spectrum. On the other hand, $\sim 40\%$ of these unobscured type-2s have $L_{2-10keV} < 2 \times 10^{43}$ erg s^{-1}, i.e. some of them can have a significant contribution from a starbursts in the soft band. Finally, some unobscured type-2s may be faint type-1 AGNs, with strong optical/IR contamination from host galaxy light, so that the SED fitting fails to reveal their intrinsic nature.

All the sources classified as galaxies have low upper limits in column density, consistent with Galactic values. This is compatible with the fact that the X-ray emission of these sources comes from the integrated emission of X-ray binaries, plus diffuse hot gas (Fabbiano 1989). The emitting region therefore is extended, not nuclear, and no global obscuration is expected, unless the galaxy is observed perfectly edge-on and the whole disk is obscured by a dust lane.

5.2 Γ, photon index

In the right panel of Fig. 8 the distribution of Γ for the CCBS is shown. The distribution of Γ for type-1 AGN is strongly peaked in the bin between 1.8 and 2, with a mean value $\langle \Gamma \rangle = 1.89 \pm 0.02$ and $\sigma_{obs} = 0.28$. For type-2 AGN the distributions has a peak at slightly lower values, $\Gamma = 1.6 - 1.8$, with a mean value $\langle \Gamma \rangle = 1.76 \pm 0.02$, and $\sigma_{obs} = 0.38$. The observed dispersion for type-2 is larger than for type-1, however type-2 have typically larger errors on Γ, due to higher values of N_H. Indeed, the intrinsic dispersion that we estimate taking into account the error distribution are very similar to each other ($\sigma_{int} = 0.11$ and 0.13 for type-1 and type-2 respectively).

A 2-sided Kolmogorov-Smirnov (KS) test gives a probability $P = 4.04 \times 10^{-5}$ that the two distributions are drawn from the same population. This difference can be due to a residual obscuration unaccounted in the fit, or to the contribution of a reflection component, that is known to increase in importance with decreasing luminosities (Iwasawa & Taniguchi 1993); indeed type-2 have on average lower luminosities than type-1 in the CCBS (see Fig. 9 right panel). The average quality of our spectra does not allow to test this hypothesis using a model that includes Compton reflection. Bianchi et al. (2009) verified, on the much higher quality spectra of the CAIXA AGN sample (mostly local) that the inclusion of a luminosity-dependent reflection component does not affect the differences between the two populations.

Among the galaxies, only one shows a very soft spectrum, compatible with thermal emission from hot gas, while the others have spectra that are indistinguishable from those of AGN for what concern the photon index. This can be the result of the super imposition of different components, as mentioned above.

5.3 Flux and Luminosity

Fig. 9 (left panel) shows the distribution of the 2-10 keV band observed flux for the CCBS sources. The flux distributions for type-1 and type-2 are very similar (KS probability $p = 0.13$) and broad, as expected from the flux limited nature of our sample, with similar
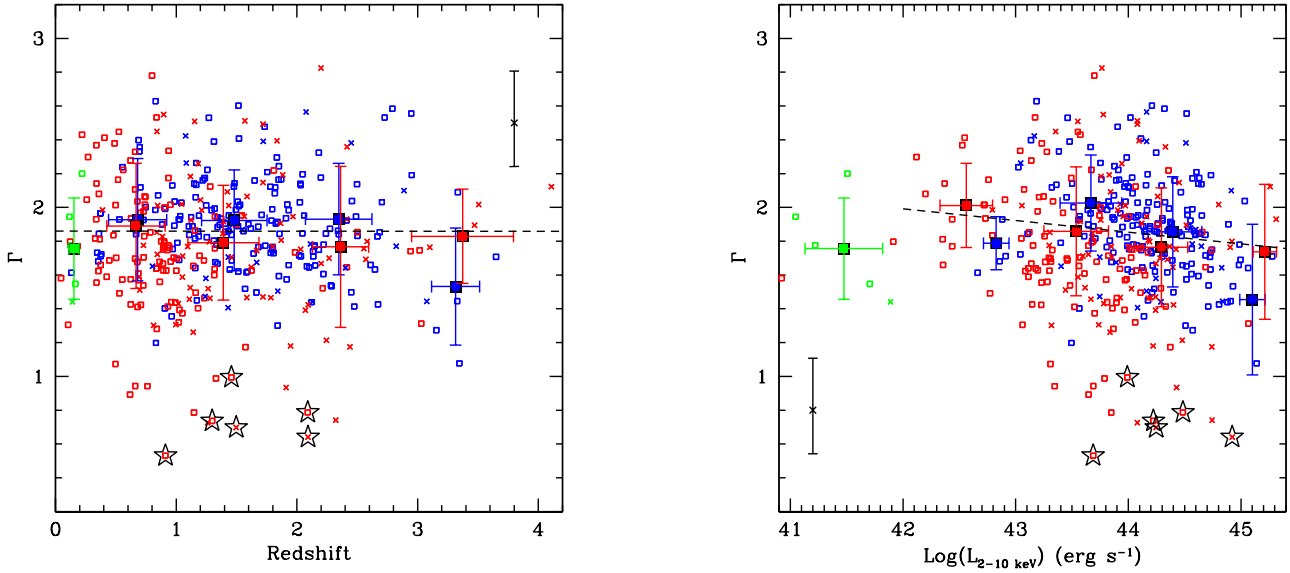
Figure 10. *Left panel:* Distribution of photon index as a function of redshift for the CCBS sources. *Right panel:* Distribution of Γ as a function of the 2-10 keV luminosity, corrected for absorption, for our sources. In both panels, small symbols represent different classes of sources as in previous figures. The large thick squares represent the weighted mean for bins of redshift ($\Delta z=1$) and $\Delta Log(L_{2-10keV}) = 1$, respectively. The error bars mark the 1σ dispersions in both axes. The dashed lines represent the least-squares fit to the type-1 and type-2 AGN together (CT candidates are excluded from the fit).

average values $\langle F_{2-10keV} \rangle \sim 1.6 \times 10^{-14}$ and standard deviation $\sigma \sim 2.4 \times 10^{-14}$ erg cm^{-2} s^{-1}.

Fig. 9 (right panel) shows the distribution of intrinsic 2-10 keV luminosity, corrected for absorption. We computed the intrinsic 2-10 keV luminosity from the best fit model and the redshift of each source, imposing the intrinsic and Galactic absorption to be zero. Type-1 AGN are typically more luminous in the 2-10 keV band, with 63% of them being in the quasar regime, i.e. $L_{2-10keV} > 10^{44}$ erg s^{-1}. Type-2 AGN are less luminous, with only 33% of the sources having $L_{2-10keV} > 10^{44}$ erg s^{-1}. The two distributions are significantly different (KS test probability $p = 1.58 \times 10^{-10}$). This difference is the result of the different redshift distributions (type-1 AGN typically have higher redshift) coupled with the fact that our sample is flux-limited, i.e. redshift and luminosities are tightly correlated (see Fig. 2).

5.4 Photon Index dependencies

We studied the possible evolution of the photon index with redshift and X-ray luminosity. Fig. 10 (left panel) shows the distribution of Γ a function of redshift, while the right panel shows the distribution of Γ as a function of the 2-10 keV intrinsic luminosity, corrected for absorption. Individual sources are color coded as in Fig. 4. The thick filled squares represent the weighted mean for bins of redshift ($\Delta z=1$) and luminosity ($\Delta Log(L_{2-10keV})=1$) respectively. The error bars mark the 1σ dispersions in both axes in the two plots. The CT candidates sources are marked with stars, and have been excluded from the least-square fit and mean.

The dashed line in both panels represents the least-squares fit, weighted by the errors in Γ, for type-1 and type-2 AGN together. We excluded from the fit sources classified as galaxies. There is no evident correlation in redshift, apart for a hint of harder spectra for type 1 AGN at $z > 3$ (the average is at $\sim 1\sigma$ from the typical value). However we stress that there are only 6 type 1 AGN at $z > 3$, therefore we do not have enough statistic to

prove that the effect is real. This could be due to the fact that at these high redshifts, the reflection hump enters the observing band (30 keV at $z = 3.5$ correspond to 6.7 keV) flattening the observed spectrum. At the average luminosity of the sources in this bin ($\langle L_{2-10keV} \rangle = 6.8 \times 10^{44}$ erg s^{-1}) the EW of the Fe Kα is in the range $EW = 30 - 60$ eV (Bianchi et al. 2007) and the contribution of the reflected component is only $\sim 10\%$ in the 2-7 keV observed band. We simulated spectra with the appropriate reflection component superimposed to the primary power-law (with $\Gamma = 1.9$) as underlying model, with ~ 1000 and ~ 100 counts (1000 spectra for each counts regime), and applied the same spectral analysis described in Sec. 4.1. The average photon index resulting from a fit with a simple power-law hare $\langle \Gamma \rangle = 1.58 \pm 0.08$ and $1.62^{+0.27}_{-0.25}$, respectively. Thus the observed hardening of the average Γ at $z > 3$ is consistent with being produced by an unaccounted reflection component. However, if this is the correct explanation for that effect, it doesn't' explain why this effect should be present for type 1s and not for type 2s. We also underline that we only have 6 type 1s and 7 type 2s at such high redshift, therefore we do not have enough statistics to draw any firm conclusion.

In the Γ vs $L_{2-10keV}$ plot a weak anti-correlation can be seen in both type-1 and type-2 classes, with a slope of -0.09 ± 0.06 (but consistent with no correlation within 1.5σ). A steeper anti–correlation between photon index and hard X-ray luminosity has been found in previous work on *Chandra* and *XMM–Newton* counterparts of SDSS quasars (Young et al. 2009, Green et al. 2009, Kelly et al. 2009). On the other hand, no correlation was found between these quantities in the deeper survey CDFS (Tozzi et al. 2006. As mentioned above, the current interpretation of this effect is that there is an increasing contribution of the reflection component, that at high redshift (and hence luminosity) enters the observing band. The decrease of the normalization of the reflection component with increasing luminosity is not strong enough to compensate this effect.

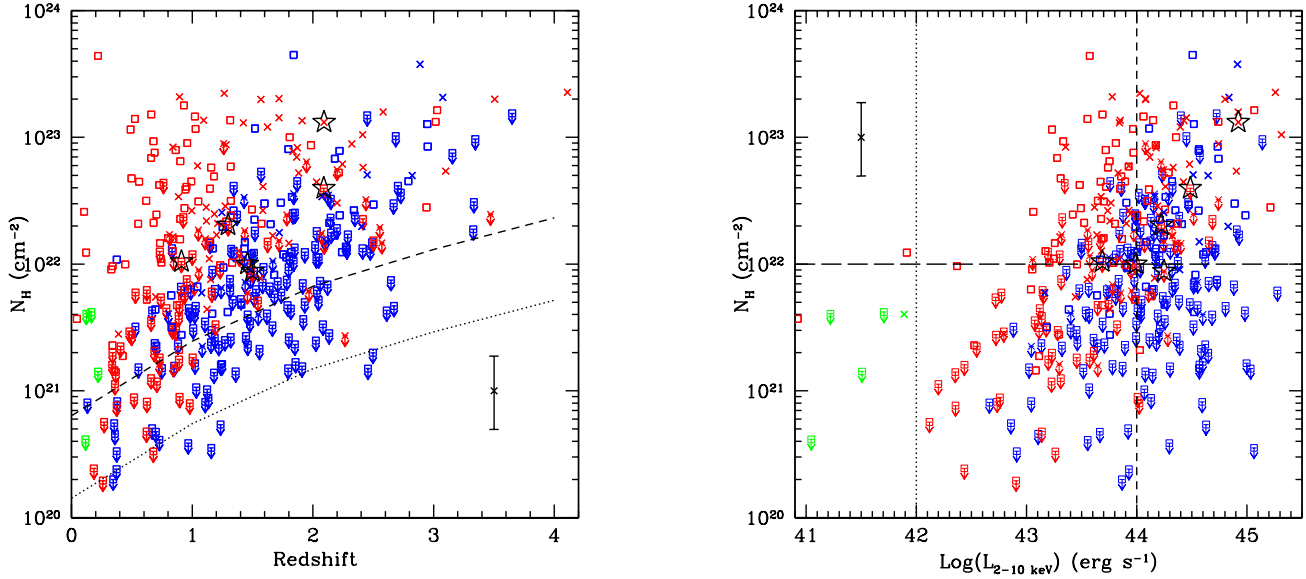Making the plot of Fig. 10 (left) in logarithmic scale instead

Figure 11. *Left panel:* Distribution of intrinsic absorption, as a function of the source redshift for the CCBS. Symbols as in previous figures. The dashed (dotted) curve represents the upper limit obtained from the spectral fit of a 200 (1000) counts simulated spectrum with no intrinsic absorption and photon index $\Gamma = 1.9$, for a given redshift. *Right panel:* Distribution of intrinsic absorption, as a function of the 2-10 keV intrinsic luminosity, corrected for absorption. Symbols as in previous figures. The dashed line marks the separation between quasars and Seyfert ($L_{2-10keV} = 10^{44}$ erg s^{-1}), the long dashed line represents the N_H value used to separate obscured and unobscured sources ($N_H = 10^{22}$ cm^{-2}) while the dotted line shows the luminosity cut we adopted to distinguish between AGN and star forming galaxies ($L_{2-10keV} = 10^{42}$ erg s^{-1}).

of linear, it would produce a distribution similar to that of the Γ vs L_{2-10} plot (right). The two quantities clearly go together in a flux/counts limited sample, but there is also a selection effect here, for which, at a given flux/counts, sources with flatter spectra have higher $L_{2-10keV}$ (right). Binning in linear z puts together sources with $L_{2-10keV}$ in 1.5-2 dex range (at z $\stackrel{<}{\sim}$ 2), thus diluting the effect.

5.5 Column density dependencies

Fig. 11 (left panel) shows the distribution of intrinsic absorption, as a function of source redshift. We note that there is a strong dependency of the minimum value of N_H (either upper limit or detection) in the sample, and the redshift. As already noted in Tozzi et al. (2006), this effect is due to the fact that, with increasing redshift, the photoelectric absorption cut-off moves toward the lower limit of the observing band (i.e. 0.5 keV), and it becomes more difficult to measure small values of N_H. This can introduce a bias in the distribution of the observed fraction of obscured sources with redshift (Gilli et al. 2010). To illustrate this effect, the dashed (dotted) curve in the plot represents the upper limit obtained from the spectral fit of a 200 (1000) counts simulated spectrum with no intrinsic absorption and photon index $\Gamma = 1.9$, at any given redshift.

Furthermore, there is a lack of sources with high column densities ($N_H > 10^{22}$ cm^{-2}) at low redshift ($z < 0.5$). This can be explained by the fact that a low luminosity, low redshift source, with a high column density would have a highly suppressed flux (the photoelectric absorption cut off is at few keV), and it would not pass our counts–based selection criterion. There are just two exceptions: one being the source CID-1105, at z=0.219. This source shows, in addition to the primary power law, that is completely suppressed below 2.5 keV, a second soft component, which contributes for $\sim 25\%$ to the total number of counts (see Sec. 4.3). The other

exception is source CID-1678 at z=0.104, that has a much lower column density and also shows a hint of a second, soft component below 1 keV, but less prominent than in CID-1105.

Fig. 11 (right panel) shows the distribution of intrinsic absorption N_H as a function of the 2-10 keV intrinsic luminosity, corrected for absorption. The dashed vertical line marks the adopted separation between quasars and Seyfert ($L_{2-10keV} = 10^{44}$ erg s^{-1}), the long dashed horizontal line represents the N_H value used to separate obscured and unobscured sources ($N_H = 10^{22}$ cm^{-2}), while the dotted vertical line shows the luminosity cut we adopted to distinguish between AGN and star forming galaxies ($L_{2-10keV} = 10^{42}$ erg s^{-1}, see Sec. 3.3).

As noted in Sec. 4.4 for the distribution of N_H and $L_{2-10keV}$ separately, type-2 AGN have higher values of N_H and lower luminosities, so that the region of $N_H > 10^{22}$ cm^{-2} and $L_{2-10keV} < 10^{44}$ erg s^{-1} is populated almost only by type-2 sources, while almost all the sources in the opposite region ($N_H < 10^{22}$ cm^{-2} and $L_{2-10keV} > 10^{44}$ erg s^{-1}) are type-1s. As mentioned for the N_H vs. z distribution, sources at low luminosities and high N_H are missed in our counts–limited sample due to selection effects.

About 15% (62 out of 390) of all the sources lie in the region of obscured luminous quasar ($N_H > 10^{22}$ cm^{-2} at 1σ and $L_{2-10keV} > 10^{44}$ erg s^{-1}). Such high values of the fraction of X-ray selected type 2 quasar have been found in deeper *Chandra* and *XMM–Newton* surveys (Mateos et al. 2005, Tozzi et al. 2006), and are compatible with the predictions of X-ray background synthesis models ($\sim 12\%$ in the flux range of the CCBS, Gilli et al. 2007). We underline here the importance of the 100% completeness in redshift of CCBS: typically not identified sources have faint optical counterparts, and these are usually obscured sources at high redshift-high $L_{2-10keV}$ (Fiore et al. 2003). Most of the sources in the type-2 quasar region are classified as galaxies at high redshift, through their SED fitting. The excellent quality of COSMOS pho-
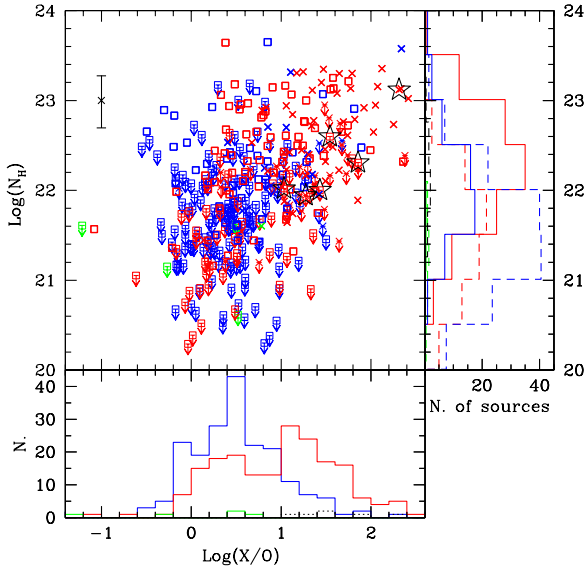
Figure 12. Distribution of X/O as a function of N_H, for the 390 sources in the CCBS. The bottom and right projections show the distribution of X/O and N_H respectively for each class of sources. Symbols and colors as in previous figures.

tometric redshifts is indeed one of the unique features of COSMOS as an omni-wavelength survey, allowing to efficiently identify this class of elusive sources.

Twenty sources classified as type-1 AGN are in the region of type-2 quasar (out of 62). These are very interesting sources, given that they are natural candidates to be Broad Absorption Lines (BAL) quasars. BAL quasars show broad, blue shifted absorption lines in the optical/UV spectra (Turnshek et al. 1980; Weymann et al. 1991). They also show strong obscuration in X-rays, typically from a 'warm' absorber (Reynolds 1997; Piconcelli et al. 2005) and/or in the form of blue shifted absorption lines of highly ionized gas (Fe XXV, Fe XXVI) (Chartas et al. 2002; Markowitz et al. 2006; Tombesi et al. 2010; Lanzuisi et al. 2012). Of these 20 BAL QSO candidates, 14 have an optical spectrum available, and almost all (13/14) of them lie at redshift $z > 1.5$, i.e. it should be possible to observe such UV features (e.g. C IV at $\lambda 1549$Å) from the ground in their optical spectra. Indeed, to a visual inspection of their optical spectra, at least 50% of them shows absorption signatures at the expected locations of ionized metals typically observed in BAL quasars, such as Mg II, Al III, Si IV, C IV, N V and O VI. The proper analysis of the optical spectra of these BAL candidates is ongoing and will be the subject of a dedicated paper.

5.6 X/O ratio

We studied the distribution of the X/O ratio, defined as in Civano et al. (2012), i.e. $X/O = log(F_X/F_{Opt})$[14]. It is well known that the X/O ratio is a z dependent quantity for obscured AGN, given that optical and X-ray bands have a k-correction going in opposite directions: the k-correction is negative in the optical band and positive for the X-rays (Comastri et al. 2003; Fiore et al. 2003; Brusa et al. 2010). As a result, obscured sources at high redshift have

---

[14] We used the 2-10 keV band and the i optical magnitude.

higher X/O. On the other hand, unobscured sources have similar k-corrections in the two bands, and the distribution in X/O is not correlated with the redshift (see Fig. 14 of Paper III).

The X/O ratio is known to be a proxy for the absorbing column density, at least in the Compton thin regime (Fiore et al. 2003). However, the actual correlation between N_H and X/O has not been demonstrated yet directly due to the lack of statistics in previous samples. Fig. 12 shows the distribution of X/O as a function of N_H, for the 390 sources in the CCBS. The right and lower projections show the distribution of N_H and X/O, respectively, for each class of sources. For type-1 AGN the distribution of X/O peaks at X/O $\simeq$ 0.5 and the majority of these sources are distributed between 0 and 1, while the observed range for optically and soft X-ray selected AGN is typically more extended toward lower values (i.e. $-1 <$ X/O < 1; Zamorani et al. 1999, Lehmann et al. 2001). This can be explained considering that we are sampling the bright end of the $F_{2-10keV}$ distribution of AGN in C-COSMOS, i.e. we are selecting the X-ray brightest sources.

Type-2 AGN peak at X/O > 1, reflecting the fact that these sources are on average more obscured in the optical band, although a tail to low X/O and low N_H is present. The median value for type-2s in our sample is $\langle X/O \rangle = 0.96$, with dispersion of $\sigma = 0.64$. For comparison, the median value and dispersion for the sample of AGNs presented in Civano et al. 2012, combining *Chandra*-COSMOS and *XMM–Newton*-COSMOS X-ray catalogs are $\langle X/O \rangle = 1.02^{+0.59}_{-0.72}$ and $\langle X/O \rangle = 0.82^{+0.66}_{-0.74}$ for type-2s and galaxy SEDs respectively.

Among these sources there is a large fraction of sources with optical/IR properties of passive galaxies, that are classified as AGN thanks to the multi-wavelength coverage and SED fitting together with X-ray informations. Indeed, the typical N_H for these sources is in the range $10^{22} - 10^{23}$ cm^{-2} and, if converted into optical extinction A_V through the standard Galactic dust-to-gas ratio ($A_V = N_H/(1.8 \times 10^{21})$, e.g. Predehl & Schmitt 1995), translates into an extinction $A_V \gtrsim 5$, large enough to completely obscure the AGN contribution in the optical bands.

The distribution of X/O vs. N_H for type-2 AGN closely resembles the one found between N_H and 2-10 keV luminosity. This is due to the fact that mildly obscured AGNs (up to few$\times 10^{22}$ cm^{-2}) have their nuclear optical/UV light strongly attenuated by dust and, therefore, the X/O ratio represents the ratio between the X-ray flux from the AGN and the host galaxy starlight, which spans a smaller luminosity range (Fiore et al. 2003).

Almost all type-1 AGN with $N_H > 10^{22}$ cm^{-2} (i.e. BAL quasar candidates), have an X/O ratio in the range 0-1, typically lower than the population of type-2 obscured sources. This is in agreement with the idea that these sources are obscured by outflowing gas in the very central regions, blocking the X-rays coming from the inner regions of the system, but not the optical emission coming from much larger regions.

Nearly 70% of all the sources with X/0> 1 have $N_H > 10^{22}$ cm^{-2}, thus confirming the idea that the X/O ratio is a proxy for the obscuring column density, and selecting sources above this threshold is an efficient tool for the creation of samples of obscured AGN (Fiore 2003; Comastri & Fiore 2004). However, there is a 30% of contamination by sources with high Log(X/O) and low X-ray obscuration, as found in Tozzi et al. (2006).

6 SUMMARY

We have performed the spectral analysis of the 390 brightest extragalactic sources in the *Chandra*-COSMOS field with more than 70 net counts in the 0.5-7 keV band. The sample has a 99% completeness in optical identification, $\sim 75\%$ of the sources have a spectroscopic redshift, while the remaining $\sim 25\%$ have a photometric redshift.

The spectral fit was performed using a non binned minimization technique and assuming as underlying model a simple power law, modified by intrinsic absorption at the redshift of the source, plus galactic absorption fixed to the value in the direction of the field. Simultaneously a global model for the background was fitted to the local background spectra, leaving 2 free parameters to account for local variations.

The main observational results can be summarized as follows:

• In appendix A we show that our spectral fit procedure, based on the use of Cstat statistic and simultaneous total and background fit, proved to be accurate and reliable: the spectral parameters obtained are stable, regardless of the combination of statistic and number of counts per bin we chose, while for the errors we proved that the Cstat applied to binned spectra (with 5 counts per bin) gives more constrained results (i.e. relative errors $< 30\%$ for most sources) with respect to a χ^2 analysis. The Cstat applied to non binned spectra tends to underestimate the errors.

• The sample is divided almost equally in type-1 (49.7%) and type-2 AGN (48.7%) plus few passive galaxies at low z. However 70% of type-2 AGN has a spectral or photometric classification of passive/starforming galaxy, or with hybrid or ambiguous classification, and an 'a posteriori' selection criterion, based on the X-ray luminosity, was required to identify them as sources hosting a hidden AGN. Indeed, at the flux level of our sample almost all the sources classified as non active turned out to be obscured AGN.

• The average photon index of type-1 AGN is $\langle \Gamma \rangle = 1.89 \pm 0.02$ with intrinsic dispersion $\sigma_{int} = 0.11$, while for type-2 AGN is $\langle \Gamma \rangle = 1.76 \pm 0.03$ and $\sigma_{int} = 0.13$. The two distributions have a high probability to be intrinsically different (KS test probability 4.04×10^{-5}). With the current data it is not possible to determine if this is an intrinsic difference or it is due to unaccounted extra components in the spectra.

• The distribution of column density for type-1 AGN is dominated by upper limits (75%), anyway $\simeq 17\%$ of them results obscured from the X-ray spectral analysis. For type-2 AGN the distribution is broader, with a peak around 10^{22} cm^{-2} and 60% of the sources with a detection of the value of the column density. Only $\sim 50\%$ of type-2 can be considered obscured. This is a low value compared with previous results. Possible explanations can be a significant contribution from starbursts emission in the soft band, or misclassification of faint type-1 with strong optical/IR contamination from host galaxy light.

• No obvious CT source is present in the CCBS, even if a sample of 6 CT candidates has been selected on the basis of their very flat X-ray spectra. For only 2 of these sources, an emission Fe line is detected, at energies slightly different, but still consistent with the expected value.

• The intrinsic 2-10 keV luminosity, corrected for absorption, of type-1 AGN peaks at $\simeq 10^{44}$ erg s^{-1} and 63% of these sources are in the quasar regime. The majority ($\sim 70\%$) of type-2 AGN instead are concentrated below this value and are therefore mostly Seyfert galaxies. This difference is clearly due to the combination of an X-ray flux limited sample and the different redshift distributions, with type-1 AGN having typically higher redshift.

• No correlation has been found between the photon index and the redshift for our sources, while a weak anti-correlation is observed in X-ray luminosity. The correlation (with slope -0.09 ± 0.06) is weaker of what has been found in previous work on *Chandra* and *XMM–Newton* counterpart of bright SDSS quasars, while consistent with no correlation as found in the CDFS.

• About 15% of the sources lie in the region of type-2 QSOs, a value in agreement with deeper X-ray surveys and with population synthesis models. This high fraction was made possible by the almost 100% completeness of our sample and the accuracy of photometric redshifts: a large fraction of these QSO2 are indeed high redshift galaxies with clean counterpart only in near and mid IR. 20 sources are type-1 BAL QSO candidates. Of them 14 have an optical spectrum and at least 50% of them indeed show indication of absorption lines in their optical spectra.

• The distribution of X/O for type-1 AGN peaks at X/O $\simeq 0.5$ and the vast majority of these sources are distributed between 0 and 1. Type-2 AGN peak at X/O $\gtrsim 1$ and 50% of them show $1 < $ X/O < 2, an extreme value for AGN. Furthermore, nearly 70% of all the sources with X/O > 1 are obscured, thus confirming that selections based on high X/O ratio are efficient in finding samples of obscured AGN.

• In appendix B we show that there is a small systematic difference in the spectral parameters obtained from *Chandra* and from *XMM–Newton* (280 *XMM–Newton* counterparts of the CCBS sources). The *XMM–Newton* spectra tend to be fitted with softer power laws (up to 20% difference) and lower obscuration (5% difference in the fraction of obscured sources) with respect to *Chandra* spectra.

The CCBS represents just the tip of the iceberg of the complete C-COSMOS catalog, and a large number of interesting sources such as heavily obscured AGN and starbursts galaxies are located below our selection threshold. The use of our well tested procedure on this larger sample will give important insights on the co-evolution of accretion and star formation.

Finally, the *Chandra* COSMOS-Legacy large program has been recently approved (Cycle 14), aimed to extend the *Chandra* coverage of the COSMOS area, taking ~ 1.45 deg^2 to 160 ks depth and covering the remaining 0.75 deg^2 field to 100-50 ks depth (for a total of 2.8 Msec). It is expected to produce a final sample of bright sources almost 3 times larger than the current one (~ 1100 sources) allowing for a much more accurate determination of the intrinsic distribution of the X-ray spectral parameters in the sample.
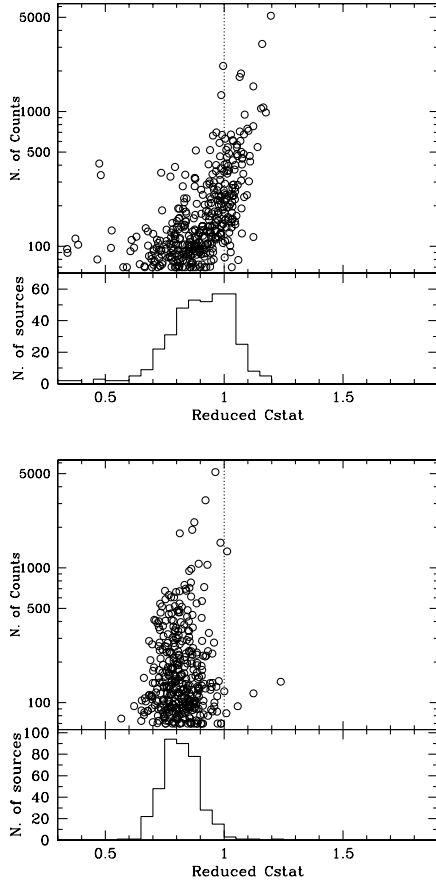
[h]

Figure A1. Distribution of $Cstat_\nu$ as a function of counts for the not binned case (top) and the 1 counts per binned case (bottom).

sion at NASA/GSFC and the High Energy Astrophysics Division of the Smithsonian Astrophysical Observatory.

APPENDIX A: CSTAT AND χ^2

We performed a series of tests to verify the behavior of the Cstat statistic in the range of counts available for our sample, from 70 to few thousands net counts. We verified that the $Cstat_\nu$, applied to unbinned data, is strongly dependent on the number of counts, moving to ~ 1 for sources with few thousands counts, to $\sim 0.7 - 0.8$ for sources with $70 - 100$ counts (Fig. A.1 top panel). This happens because of the high number of empty channels for the sources with few tens of counts, which contribute to the total number of DOF, but not to the amount of Cstat. On the other hand, applying a binning of 1 counts per bin removes the dependency of $Cstat_\nu$ from the number of counts, because the empty channels are avoided. However, the distribution of $Cstat_\nu$ is not centred at ~ 1 as expected (and as for the 5 binning case) but instead at lower values (~ 0.85, Fig. A.1 bottom panel).

Interestingly, the resulting spectral parameters obtained in the three ways are almost indistinguishable. Also the associated errors are comparable, even if, when no counts binning is applied, there is a population of sources for which the errors are typically smaller, with respect to both fits with 1 and 5 counts per bin. Fig. A.2 (panel
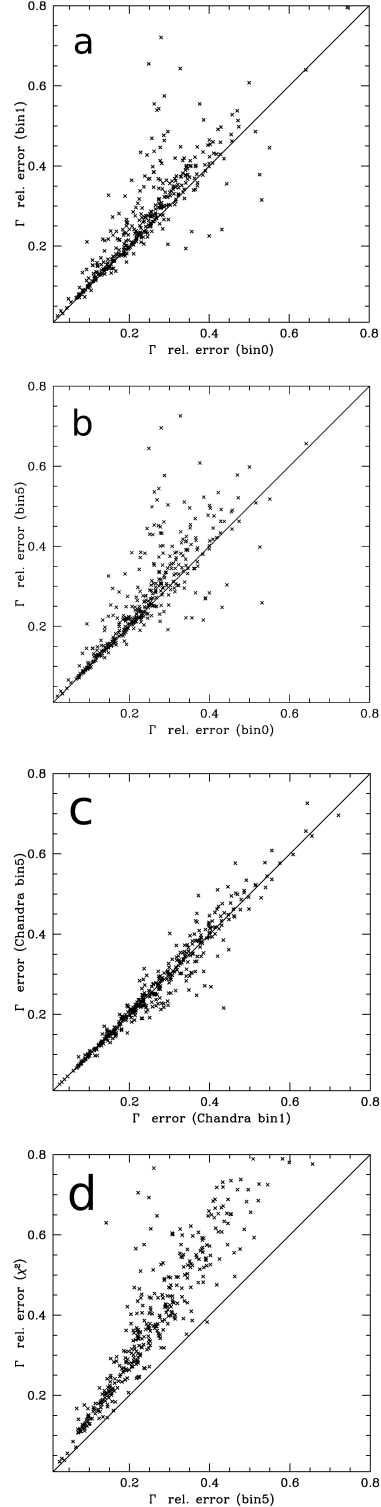

Figure A2. Comparison between the relative errors on Γ for different binning and statistics adopted: Cstat(bin0) vs. Cstat(bin1); Cstat(bin0) vs. Cstat(bin5); Cstat(bin1) vs. Cstat(bin5) and Cstat(bin5) vs. χ^2(bin15), from top to right).

a) shows the comparison between the relative errors on Γ measured from Cstat with no binning and from Cstat with 1 counts per bin. Panel b shows the comparison between Cstat with no binning and Cstat with 5 counts per bin. Not surprisingly, the average difference between errors increases with higher values of relative errors, i.e. lower counts available in the spectra. Panel c shows the comparison between Cstat with 1 counts per bin and Cstat with 5 counts per bin, that are comparable.

To estimate the expected errors on the spectral parameters, we simulated 1000 spectra in four counts regime, i.e. 100, 200 500 and 1000 net counts, respectively. The input model has fixed values of $\Gamma = 1.9$, $N_H = 1 \times 10^{22}$ cm^{-2} and $z = 1.0$ (i.e. close to the average value for our sample). We then performed our analysis with the Cstat and the three different binning (bin0, bin1 and bin5). The 1σ dispersions on Γ are comparable in the three cases, and of the order of 0.25, 0.19, 0.15 and 0.09, that translate into relative errors of 13,10,8 and 5% respectively, thus fully in agreement with the distribution of relative error show in Fig. 5 (right panel).

We also performed a fit applying the χ^2 statistic for all the sources, after binning the spectra to 15 counts per bin. The best fit spectral parameters are again very similar to those derived from the Cstat statistic (regardless of the number of counts per bin). Instead, the average ratio between the relative errors obtained with Cstat (5 counts per bin) and with the χ^2 is $Err_{Cstat}/Err_{Chi2} = 0.69$, with a small dispersion ($\sigma = 0.11$), i.e. the Cstat gives error typically $\sim 30\%$ smaller than χ^2, for our sources (see also Tozzi et al. 2006), with increasing differences at decreasing number of counts, enforcing our choice to use the Cstat instead of χ^2 statistic. In Fig. A.2 (d panel) is reported the comparison between the relative errors on Γ computed from Cstat with 5 counts per bin and from χ^2 with 15 counts per bin.

APPENDIX B: COMPARISON WITH XMM-NEWTON COSMOS DATA

The COSMOS field has been observed in X-rays by *XMM–Newton* between 2003 and 2006. The *XMM–Newton* observation covers a region of 1.4×1.4 deg^2, to a sensitivity limit of 0.7×10^{-15} erg cm^{-2} s^{-1} in the 0.5-2 keV band (Hasinger et al. 2007). Details on the point-source detection method for the XMM-COSMOS survey are presented in Cappelluti et al. (2007), while details on extraction of the X-ray spectral counts for the sources in the XMM-COSMOS catalog, are described in Mainieri et al. (2007). As for C-COSMOS sources, the spectral counts were extracted separately from each observations and merged together, weighting the response matrices ARF and RMF for the contribution of each spectrum. We applied the same cut in number of counts (70 net counts in 0.5-7 keV) to collect a sample of 280 *XMM–Newton* counterparts of the CCBS sources.

The spectral analysis was performed in the same way described in Sec 4, i.e. using the Cstat minimization technique, applying a 5 counts binning and simultaneously fitting the background. The energy band taken into account was limited to the 0.5-7 keV band, equal to the one used for the fit of *Chandra* data, even if the nominal sensitivity of *XMM–Newton* reaches ~ 10 keV, in order to minimize systematic effects due to the different bandpass.

B1 Photon index

Fig. B.1 (left panel) shows the normalized difference in photon index (($\Gamma_{XMM}/\Gamma_{Chandra}$) $- 1$, i.e. $\Delta\Gamma/\Gamma = 0.1$ is equivalent to a

10% difference in Γ) as a function of the number of *XMM–Newton* counts for all the 280 *XMM–Newton* counterparts of the CCBS sources. The red squares represent the median of the distribution of $\Delta\Gamma/\Gamma$, computed in bins of counts (70-150; 150-300; 300-600; 600-1000 and >1000). We note that there is a general good agreement between the two values, but for bright sources (i.e. with more than $\sim$ 1000 counts) the distribution is asymmetric and the difference tend to be always positive, i.e. the *XMM–Newton* photon index is typically higher than the *Chandra* one, with the difference spanning the range 0-20%. Furthermore for most of these sources the relative difference is not consistent with zero even taking in to account the error bars. This trend is present, although less significantly because of the larger errors, in the bins with a small number of counts.

Given this unexpected result, we decided to further investigate this issue. We collected the photon indices computed in Mainieri et al. (2007) for a sample of 135 AGN in XMM-COSMOS survey, showing more than 100 total counts and having a spectroscopic identification. The spectral fit of these sources was performed in a completely different way, using the χ^2 minimization technique, and performing background subtraction and counts binning. 58 of these sources are in common with our sample. Fig. B.1 (right panel) shows the relative difference between the photon index obtained from our analysis of *XMM–Newton* data and the one independently obtained in Mainineri et al. (2007), as a function of *XMM–Newton* counts for these sources. These values are in very good agreement, at all counts level, and in any case consistent within the error bars for almost all the sources. This analysis suggests that we are observing an unexpected systematic difference in the determination of the photon index of the order of $\sim 10\%$ from *Chandra* and *XMM–Newton*, with *XMM–Newton* spectra being softer than *Chandra* ones.

In our analysis of *Chandra* spectra, we used as calibration files CALDB 4.1.2 for the extraction of *Chandra* spectra and matrices. It is known that a major correction in the calibration files occurred in Dec. 2010, due to the contamination model, which under-predicted the attenuation by the Ca-K edge, starting from 2009[15]. A similar, but less severe, correction occurred in Dec. 2009, regarding observations taken from 2006 (C-COSMOS observations were taken mostly during 2007). As a check we extracted, with the newest calibration files available (CALDB 4.3), the spectra and matrices of the 20 brightest sources (> 500 counts) in the sample. Then we performed again the same spectral analysis on the new spectra, to verify if this new version of the calibrations has some effect on the determination of spectral parameters. The results obtained with this new calibration are completely consistent with the ones described above.

B2 Column density

Fig. B.2 (left panel) shows the comparison between the columns density obtained from the fit of *Chandra* data and the N_H obtained from *XMM–Newton* for the sample of 280 sources in the CCBS with *XMM–Newton* counterpart. Magenta points represent N_H detections from both data sets, green (blue) points represent *Chandra* (*XMM–Newton*) upper limits and *XMM–Newton* (*Chandra*) detection of the N_H, and red points represent upper limits on N_H from both datasets The thick line represents the 1:1 relation,
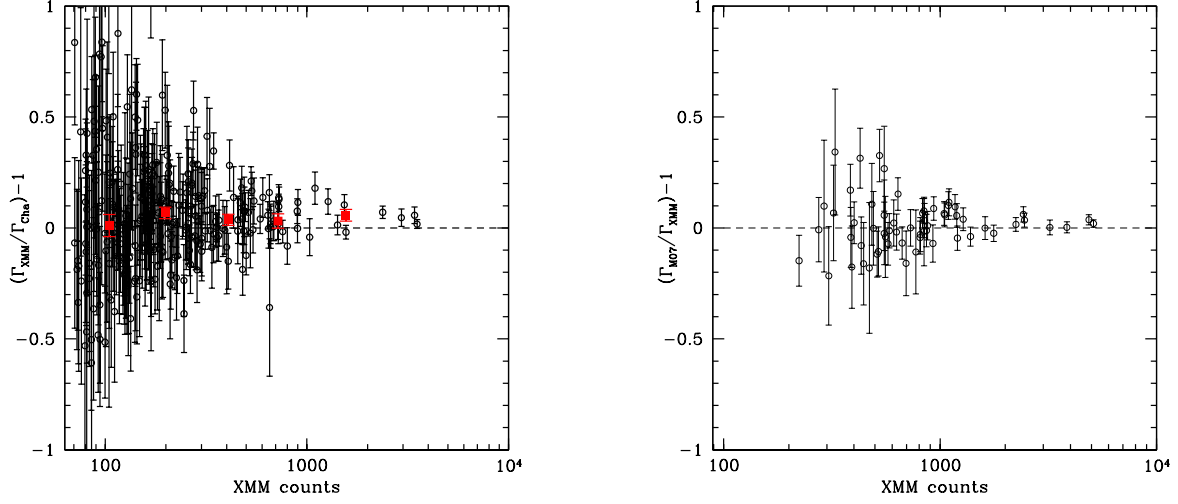
---

[15] http://cxc.cfa.harvard.edu/ciao/why/acisqecontam.html

Figure B1. *Left panel:* Normalized difference in photon index $((\Gamma_{XMM}/\Gamma_{Chandra}) - 1)$ as a function of the number of *XMM–Newton* counts for all the 280 sources with *XMM–Newton* counterpart showing more than 70 counts. The red squares represent the median of the distribution of $\Delta\Gamma/\Gamma$ computed in bins of counts (70-150; 150-300; 300-600; 600-1000 and >1000). The vertical bars show the standard error on the median for each bin. *Right panel:* Normalized difference between the photon index obtained from our analysis of *XMM–Newton* data and the one independently obtained in Mainieri et al. 2007, for the 58 sources in common.
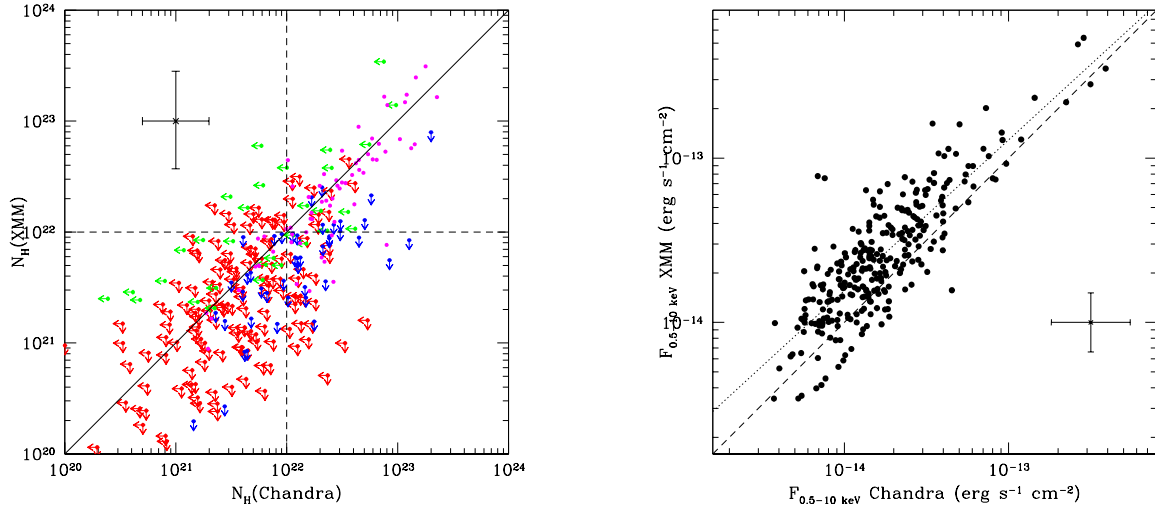

Figure B2. *Left panel:* Comparison between the columns density obtained from the fit of *Chandra* data, and the N_H obtained from *XMM–Newton* for the 280 counterparts of the CCBS. Different colors represent different combinations of detection and upper limits in the measured value of N_H (see text). The thick line represents the 1:1 relation, the dashed lines mark the separation between obscured and unobscured sources, in both axes. *Right panel:* Comparison between the 0.5-10 keV band fluxes obtained from the two sets of data for the sample of 280 sources in common in the two surveys. The dashed line represents the 1:1 relation, the dotted line is the linear regression obtained for the data.

and the dotted lines mark the threshold assumed to distinguish between X-ray obscured and unobscured sources ($N_H=10^{22}$ cm^{-2}). For sources with a detection on N_H from both satellites, there is a good agreement in the values of N_H along all the range of values. We note that there is a small population (6 sources) that occupies the lower right quadrant (i.e. *Chandra* measures a column density $N_H > 10^{22}$ cm^{-2} while the value measured by *XMM–Newton* is lower) while the opposite region (upper left) is empty. Similarly, there are slightly more sources which result obscured from *Chandra* and for which *XMM–Newton* measures an upper limit (15), than the opposite (11 sources). This translates in a small difference on the global fraction of obscured sources in the sample, that drops

from $\sim 40\%$ to $\sim 30\%$ considering *Chandra* or *XMM–Newton* results respectively. This could in principle partly account for the mismatch historically observed between data from the Msec-deep *Chandra* fields, namely CDF-S and CDF-N and results from *XMM–Newton* surveys, at fluxes where these samples overlap (e.g. Fig. 16 of Gilli et al. 2007).

B3   Flux

Fig. B.2 (right panel) shows the comparison between the 0.5-10 keV band fluxes obtained from the two data sets for the sample of

280 sources in common. The dashed line represents the 1:1 relation, while the dotted line has been obtained as the linear regression of the two set of values. The average off-set increase with decreasing fluxes, being $\sim$ 20% at $F_{0.5-10} = 10^{-13}$ erg s^{-1} cm^{-2} and increasing to $\sim$ 40% at $F_{0.5-10} = 10^{-14}$ erg s^{-1} cm^{-2}. However, for single sources, the values for the two data sets are consistent within the error bars, for most of the sources with $F_{0.5-10} < 7-8 \times 10^{-14}$ erg s^{-1} cm^{-2} while sources with brighter fluxes have much smaller error bars, except for source CID-1678, that is at the very edge of the *Chandra* pointings, and tent to be not consistent with the 1:1 relation. The general trend can be partly due to a selection effect: we start from the *Chandra* sample, and then apply the counts cut to the *XMM–Newton* catalog. Thus we tend to select select sources with a positive fluctuation in flux in the *XMM–Newton* final sample. Finally, a few % of the sources shows a difference in the fluxes up to a factor 10, clearly indicating a real, intrinsic variability. However, the study of the X-ray variability, between the *Chandra* and *XMM–Newton* data, and inside each set of data and its relation with the optical variability (see Salvato et al. 2012, Lanzuisi et al. 2012 in preparation) will be the subject of a dedicated paper.

REFERENCES

Alexander, D. M., Aussel, H., Bauer, F. E., et al. 2002, ApJl, 568, L85

Antonucci, R. 1993, ARA&A, 31, 473

Ballantyne, D. R., Everett, J. E., & Murray, N. 2006, ApJ, 639, 740

Bianchi, S., Guainazzi, M., Matt, G., & Fonseca Bonilla, N. 2007, A&A, 467, L19

Bianchi, S., Guainazzi, M., Matt, G., Fonseca Bonilla, N., & Ponti, G. 2009, A&A, 495, 421

Blackburn, J. K. 1995, Astronomical Data Analysis Software and Systems IV, 77, 367

Bongiorno, A., Zamorani, G., Gavignaud, I., et al. 2007, A&A, 472, 443

Bongiorno, A., Mignoli, M., Zamorani, G., et al. 2010, A&A, 510, A56

Brusa, M., Comastri, A., Mignoli, M., et al. 2003, A&A, 409, 65

Brusa, M., Civano, F., Comastri, A., et al. 2010, ApJ, 716, 348

Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJs, 172, 99

Cappelluti, N., Hasinger, G., Brusa, M., et al. 2007, ApJs, 172, 341

Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635

Cash, W. 1979, ApJ, 228, 939

Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Natur., 460, 213

Chartas, G., Brandt, W. N., Gallagher, S. C., & Garmire, G. P. 2002, ApJ, 579, 169

Civano, F., Mignoli, M., Comastri, A., et al. 2007, A&A, 476, 1223

Civano, F., Elvis, M., Lanzuisi, G., et al. 2010, ApJ, 717, 209

Civano, F., Elvis, M., Brusa, M., et al. 2012, arXiv:1205.5030

Comastri, A., Mignoli, M., Ciliegi, P., et al. 2002, ApJ, 571, 771

Comastri, A., & Fiore, F. 2004, Ap&SS, 294, 63

Corral, A., Della Ceca, R., Caccianiga, A., et al. 2011, A&A, 530, A42

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156

Della Ceca, R., Caccianiga, A., Severgnini, P., et al. 2008, A&A, 487, 119

Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142

Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJs, 184, 158

Fabbiano, G. 1989, ARA&A, 27, 87

Ferrarese, L., & Merritt, D. 2000, ApJl, 539, L9

Feruglio, C., Aussel, H., Le Floc'h, E., et al. 2010, ApJ, 721, 607

Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79

Fiore, F., Grazian, A., Santini, P., et al. 2008, ApJ, 672, 94

Fiore, F., Puccetti, S., Brusa, M., et al. 2009, ApJ, 693, 447

Fiore, F., Puccetti, S., Grazian, A., et al. 2012, A&A, 537, A16

Freeman, P., Doe, S., & Siemiginowska, A. 2001, SPIE, 4477, 76

Fruscione, A., et al. 2006, SPIE, 6270

Garmire, G. P., Bautz, M. W., Ford, P. G., Nousek, J. A., & Ricker, G. R., Jr. 2003, SPIE, 4851, 28

George, I. M., & Fabian, A. C. 1991, MNRAS, 249, 352

Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79

Gilli, R. XXXXXX 2010, AIPC, 1248, 359

Green, P. J., Aldcroft, T. L., Richards, G. T., et al. 2009, ApJ, 690, 644

Haardt, F., & Maraschi, L. 1991, ApJl, 380, L51

Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417

Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, ApJs, 172, 29

Ho, L. C., Filippenko, A. V., & Sargent, W. L. 1995, ApJs, 98, 477

Hopkins, P. F., Cox, T. J., Kereš, D., & Hernquist, L. 2008, ApJs, 175, 390

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Iwasawa, K., & Taniguchi, Y. 1993, ApJl, 413, L15

Iwasawa, K., Mainieri, V., Brusa, M., et al. 2012, A&A, 537, A86

Jahnke, K., & Macciò, A. V. 2011, ApJ, 734, 92

Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775

Kelly, B. C., Bechtold, J., Siemiginowska, A., Aldcroft, T., & Sobolewska, M. 2007, ApJ, 657, 116

Kim, D.-W., Cameron, R. A., Drake, J. J., et al. 2004, ApJs, 150, 19

Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581

La Franca, F., Fiore, F., Comastri, A., et al. 2005, ApJ, 635, 864

Lanzuisi, G., Piconcelli, E., Fiore, F., et al. 2009, A&A, 498, 67

Lanzuisi, G., Giustini, M., Cappi, M., et al. 2012, A&A, 544, A2

Lawrence, A. 1991, MNRAS, 252, 586

Lawrence, A., & Elvis, M. 1982, ApJ, 256, 410

Lehmann, I., Hasinger, G., Schmidt, M., et al. 2001, A&A, 371, 833

Li, Y., Hernquist, L., Robertson, B., et al. 2007, ApJ, 665, 187

Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJs, 184, 218

Lusso, E., Comastri, A., Simmons, B. D., et al. 2012, MNRAS, 425, 623

Maccacaro, T., Gioia, I. M., Wolter, A., Zamorani, G., & Stocke, J. T. 1988, ApJ, 326, 680

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Mainieri, V., Hasinger, G., Cappelluti, N., et al. 2007, ApJs, 172, 368

Maiolino, R., Shemmer, O., Imanishi, M., et al. 2007, A&A, 468, 979

Markevitch, M., Bautz, M. W., Biller, B., et al. 2003, ApJ, 583, 70

Markowitz, A., Reeves, J. N., & Braito, V. 2006, ApJ, 646, 783

Mateos, S., Barcons, X., Carrera, F. J., et al. 2005, A&A, 444, 79

Mateos, S., Carrera, F. J., Page, M. J., et al. 2010, A&A, 510, A35

Matt, G., Perola, G. C., & Piro, L. 1991, A&A, 247, 25

McCracken, H. J., Capak, P., Salvato, M., et al. 2010, ApJ, 708, 202

Merloni, A., Bongiorno, A., Bolzonella, M., et al. 2010, ApJ, 708, 137

Mobasher, B., Capak, P., Scoville, N. Z., et al. 2007, ApJs, 172, 117

Nandra, K., & Pounds, K. A. 1994, MNRAS, 268, 405

Page, K. L., Reeves, J. N., O'Brien, P. T., & Turner, M. J. L. 2005, MNRAS, 364, 195

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193

Piconcelli, E., Jimenez-Bailón, E., Guainazzi, M., et al. 2005, A&A, 432, 15

Predehl, P., & Schmitt, J. H. M. M. 1995, A&A, 293, 889

Puccetti, S., Vignali, C., Cappelluti, N., et al. 2009, ApJs, 185, 586

Ranalli, P., Comastri, A., & Setti, G. 2003, A&A, 399, 39

Reeves, J. N., & Turner, M. J. L. 2000, MNRAS, 316, 234

Reynolds, C. S. 1997, MNRAS, 286, 513

Risaliti, G., Maiolino, R., & Salvati, M. 1999, ApJ, 522, 157

Risaliti, G., & Elvis, M. 2004, Supermassive Black Holes in the Distant Universe, 308, 187

Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, ApJ, 690, 1250

Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, ApJ, 742, 61

Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJs, 172, 86

Schinnerer, E., Smolčić, V., Carilli, C. L., et al. 2007, ApJs, 172, 46

Scott, A. E., Stewart, G. C., Mateos, S., et al. 2011, MNRAS, 417, 992

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJs, 172, 1

Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, ApJs, 172, 38

Serjeant, S., Bertoldi, F., Blain, A. W., et al. 2010, A&A, 518, L7

Siemiginowska, A., Stawarz, Ł., Cheung, C. C., et al. 2007, ApJ, 657, 145

Silk, J., & Rees, M. J. 1998, A&A, 331, L1

Silverman, J. D., Green, P. J., Barkhouse, W. A., et al. 2005, ApJ, 618, 123

Silverman, J. D., Lamareille, F., Maier, C., et al. 2009, ApJ, 696, 396

Simpson, C. 2005, MNRAS, 360, 565

Sobolewska, M. A., & Papadakis, I. E. 2009, MNRAS, 399, 1597

Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJs, 172, 9

Tombesi, F., Cappi, M., Reeves, J. N., & Braito, V. 2012, MNRAS, L413

Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, A&A, 451, 457

Treister, E., & Urry, C. M. 2006, ApJl, 652, L79

Treister, E., Urry, C. M., & Virani, S. 2009, ApJ, 696, 110

Trouille, L., Barger, A. J., & Tremonti, C. 2011, ApJ, 742, 46

Trump, J. R., Impey, C. D., McCarthy, P. J., et al. 2007, ApJs, 172, 383

Trump, J. R., Impey, C. D., Elvis, M., et al. 2009, ApJ, 696, 1195

Turnshek, D. A., Weymann, R. J., Liebert, J. W., Williams, R. E., & Strittmatter, P. A. 1980, ApJ, 238, 488

Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886

Weisskopf, M. C., Brinkman, B., Canizares, C., et al. 2002, PASP, 114, 1

Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, ApJ, 373, 23

Wright, E. L. 2006, PASP, 118, 1711

Wright, N. J., Drake, J. J., & Civano, F. 2010, ApJ, 725, 480

Young, M., Elvis, M., & Risaliti, G. 2009, ApJs, 183, 17

Zamojski, M. A., Schiminovich, D., Rich, R. M., et al. 2007, ApJs, 172, 468

Zamorani, G., Mignoli, M., Hasinger, G., et al. 1999, A&A, 346, 731